

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 15:17:31 ; Search time 109 Seconds
(without alignments)
5865.169 Million cell updates/sec

Title: US-09-976-782-15
Perfect score: 1152
Sequence: 1 cgcattggagccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq:*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq:*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	729.8	63.4	2226	2	US-08-031-538-1
2	484.4	42.0	1073	1	US-08-356-405-8
3	424.6	36.9	1686	1	US-08-356-405-1
4	416	36.1	2155	2	US-08-031-538-3
5	104.4	9.1	1382	4	US-09-016-434-1256
6	95.2	8.3	1984	4	US-09-016-434-1077
7	94	8.2	1756	4	US-09-016-434-1197
8	93	8.1	2428	3	US-08-475-742-15
9	93	8.1	2428	3	US-08-261-293-15
10	93	8.1	2818	4	US-09-376-594-629
11	91.2	7.9	1567	1	US-08-722-001-24
12	89.8	7.8	1227	1	US-08-351-473B-7
13	88.8	7.7	1738	1	US-08-334-698-3
14	88.8	7.7	1738	1	US-08-228-932-3
15	88.8	7.7	1738	1	US-08-468-939-3
16	88.8	7.7	1738	2	US-08-406-855A-3
17	88.8	7.7	1738	2	US-08-722-190-3
18	88.8	7.7	1738	3	US-09-206-899-3
19	88.8	7.7	1738	3	US-08-244-354-3
20	88.8	7.7	1738	4	US-09-444-783-3
21	88.8	7.7	1738	4	US-09-688-415-3
22	88.8	7.7	1738	4	US-09-016-434-1402
23	88.8	7.7	1738	4	US-09-444-783-3
24	88.8	7.7	1738	5	PCT-US95-04203-3
25	88.8	7.7	2108	3	US-09-032-742-6
26	88.2	7.7	1134	1	US-08-087-772A-14
27	88.2	7.7	1227	1	US-07-916-901-1

28	88.2	7.7	1270	4	US-09-016-434-1184	Sequence 1184, Ap
29	88.2	7.7	3683	3	US-08-450-962-1	Sequence 1, Appli
30	88.2	7.7	3683	4	US-08-848-631-1	Sequence 1, Appli
31	85.6	7.4	1579	4	US-09-016-434-1243	Sequence 1243, Ap
32	84.6	7.3	1987	1	US-08-722-001-26	Sequence 26, Appl
33	84.6	7.3	1997	1	US-08-722-001-27	Sequence 27, Appl
34	84.6	7.3	2004	1	US-08-722-001-11	Sequence 11, Appl
35	83	7.2	1639	1	US-08-334-698-5	Sequence 5, Appli
36	83	7.2	1639	1	US-08-228-932-5	Sequence 5, Appli
37	83	7.2	1639	1	US-08-468-939-5	Sequence 5, Appli
38	83	7.2	1639	2	US-08-406-855A-5	Sequence 5, Appli
39	83	7.2	1639	2	US-08-722-190-5	Sequence 5, Appli
40	83	7.2	1639	3	US-08-244-354-5	Sequence 5, Appli
41	83	7.2	1639	3	US-09-206-899-5	Sequence 5, Appli
42	83	7.2	1639	4	US-09-688-415-5	Sequence 5, Appli
43	83	7.2	1639	4	US-09-444-783-5	Sequence 5, Appli
44	83	7.2	1639	5	PCT-US95-04203-5	Sequence 5, Appli
45	83	7.2	1639	5		

ALIGNMENTS

RESULT 1
US-08-031-538-1
; Sequence 1, Application US/08031538
; Patent No. 5968817
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, J Gregor
; APPLICANT: Erlander, Mark G
; APPLICANT: Lovenberg, Timothy W
; TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
; TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/031,538
; FILING DATE: 19930315
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: TSR5099P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2226 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-031-538-1

Query Match 63.4%; Score 729.8; DB 2; Length 2226;
Best Local Similarity 79.6%; Pred. No. 1.3e-145;
Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;

KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;
 KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;
 KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;
 KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;
 KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;
 KW neurogenic; neuroprotective; cardiant; immunosuppressive; anorectic;
 KW virucide; receptor.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT Peptide 1..65
 FT Domain 48..68 /label= Signal_peptide
 FT Protein 66..370 /note= "Transmembrane domain"
 FT Domain 298..316 /label= Human_mature_GCREC-15
 FT Domain /note= "Transmembrane domain"

XX W0200210387-A2.

XX 07-FEB-2002.

XX 25-JUL-2001; 2001WO-US023433.

XX 27-JUL-2000; 2000US-0221478P.

XX 03-AUG-2000; 2000US-0223268P.

XX 21-AUG-2000; 2000US-0227054P.

XX 08-SEP-2000; 2000US-0231121P.

XX 13-SEP-2000; 2000US-0232243P.

XX 15-SEP-2000; 2000US-0232691P.

XX 22-SEP-2000; 2000US-0235146P.

XX (INCY-) INCYTE GENOMICS INC.

XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;
 PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Walla NK, Hafalia AJA;
 PI Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Graul RC;
 PI Warren BA, Lee EA, Ding L;

XX WPI; 2002-188744/24.

XX N-PSDB; AAD29681.

XX New human G-protein coupled receptor polypeptide for diagnosis,
 PT prevention and treatment of cell proliferative, neurological,
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic
 disorders.

XX Claim 1; Page 133-134; 150pp; English.

XX The invention relates to novel human G-protein coupled receptors (GCREC)
 CC and their encoding polynucleotides. GCREC is useful as an immunogen for
 CC preparing monoclonal and polyclonal antibodies. GCREC is useful for
 CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,
 CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,
 CC Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's
 CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,
 CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,
 CC cirrhosis, Crohn's disease), an autoimmune/ inflammatory disorder (e.g.,
 CC acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,
 CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,
 CC osteoporosis), and viral infections. GCREC is useful in a number of drug
 CC screening techniques, and to analyse the proteome of a tissue or cell
 CC type. GCREC is useful for creating knockin humanised animals or
 CC transgenic animals to model human diseases, in somatic or germline gene
 CC therapy, to generate a transcript image of a tissue or cell type, for
 CC detecting differences in the chromosomal location due to translocation,
 CC inversion, etc., among normal, carrier or affected individuals, and as
 CC hybridization probes for mapping naturally occurring genomic sequences.
 CC GCREC is useful in Southern or northern analysis, dot blot or other
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,
 CC multi-format enzyme linked immunosorbent (ELISA)-like assays, and in

CC microarrays utilising fluids or tissues from patients to detect altered
 CC GCREC expression. The present sequence is human GCREC-15
 XX Sequence 370 AA;

Alignment Scores:
 Pred. No.: 1.86e-136 Length: 370
 Score: 1861.50 Matches: 370
 Percent Similarity: 97.63% Conservative: 0
 Best Local Similarity: 97.63% Mismatches: 0
 Query Match: 86.42% Indels: 9
 DB: 5 Gaps: 3

US-09-976-782-15 (1-1152) x AAEL8654 (1-370)

QY 5 ATGAGGCGCTAGCCTTTCAGTGGCCACCGCGCGGTGGCTTGCCTGGAGCCGAG 64
 Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
 QY 65 ACCAGCAGCGGACCCGAGCCCAAGCCGAGAGGATACCTCGGTTCGACCCCGAGCGC 124
 Db 21 ThrSerSer-----GlyThrProSerProArgGlyLeuLeuGlySerThrProSerGly 38
 QY 125 GCCGTCTCTGCGGCGCGAGGCGCCCTTCTCTCTTCTTCAAGTCTCGGTGAGCGTG 184
 Db 39 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 58
 QY 185 CTAGTGTCTGATCGCTGCGACCTTCTCTGTGGAAACCTGTGTCGGTCCGATCACC 244
 Db 59 LeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro 78
 QY 245 CGGTCTCGTGTCTTCCACCGCGTCCGCAATACCTTGTGTGGCTCGACGGCGTCTCGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAATAGTGGCAGCGCTGGCGATGCCACCGAGCCTGGCGAGTGTGTGACCGGGCGA 364
 Db 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
 QY 365 CGTGGCTGTGGCGCGGAGCGCTGTGCGACGTGTGTGATCTCTTTCGACGCGCGAGCCTGT 424
 Db 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAla----- 135
 QY 425 CTGTGTCTCCCGCGCGCTCGGAAAGTGGCGGCATCGCCCTGGCGCGAGCGGGCC 484
 Db 136 LeuCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 155
 QY 485 ATCACGCGCACCTTCAGCACACGCTCGCACCCCGACGCGCGCTCGTCTCATGATC 544
 Db 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 175
 QY 545 GCGCTCCCGGCGGTGCGTGGCGCTCATGCGCTCGCGCGCGTGTCTTTTGGCGGGCG 604
 Db 176 AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 195
 QY 605 GAGGTGCGCAGCGCTCGGCTCCAGCGCTGCAGGTGAGCGGGAGACCTCTATGCGCGC 664
 Db 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 215
 QY 665 TTCTCCACCGCGCGCGCTTCCACCTCGCGCTGGCGTGGCTGGCTGTCTTACCGGAAG 724
 Db 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys 235
 QY 725 ATTCACAGCGCGCAAGTTGCTTTCGCGCGCGCGAGAGTGTGTGCGCGTCCGCG 784
 Db 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
 QY 785 GCCACCATGAGGTGAGGTCCAGGTAAAGAGACACCTGATGAGGCTGAGTGTTC 844
 Db 256 AlaThrMetGln-----ValLysGluAlaProAspGluAlaGluValValPhe 271
 QY 845 ACGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGCGGGAGCAG 904

Db 272 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTyrArgGluGln 291
Qy 905 AAGGAGAGCGGAGCGGAGGATGATGCGGGAATCTGATTGGCGCTTTGCTGCTGCTGG 964
Db 292 LysGluArgArgAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
Qy 965 ATCCCTCTTCTCTGACGGAACATCATGAGCCCACTGTGCTGCGAGCTGCCCCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
Qy 1025 TGAAGAAGCATATTTCTGTGGCTTGGCTACTCAATCTCTTCAACCCCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPheAsnProLeuIleTyr 351
Qy 1085 ACAGCTTTTAAACAAGACTAGACATGCTTCAAGAGCTCTTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 3
AAM47211
ID AAM47211 standard; protein; 370 AA.
AC AAM47211;
XX

12-FEB-2002 (first entry)

Human NOV5a protein.

Human; NOV5a; metabolic disorder; neurodegenerative disorder;
immune disorder; haematopoietic disorder; developmental disease; cancer;
retinal disease; feeding disorder; vaccine; infection; gene therapy;
neurological disorder; psychotic disorder; G-protein coupled receptor;
cytostatic; antidiabetic; virucide; neuroprotective; nootropic;
analgesic; antidepressant; antimigraine; anticonvulsant; neuroleptic;
antiallergic; antitubercular; antineoplastic; anorectic; antiarthritic;
antipsoriatic; antithrombotic; antibacterial; fungicide;
osteopathic; protozoacide; antulcer; hypertensive; hypotensive;
antinfertility; vulnerary; nephrotropic; antilipemic;
serotonin receptor; chromosome 2.

Homo sapiens.

Key Location/Qualifiers
FH Peptide 1..24
FT /label= signal_peptide
FT Protein 25..370
FT /label= mature_NOV5a

WO200174851-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US010039.

30-MAR-2000; 2000US-0193205P.

30-MAR-2000; 2000US-0193339P.

05-APR-2000; 2000US-0195343P.

06-APR-2000; 2000US-0195005P.

06-APR-2000; 2000US-0195088P.

10-APR-2000; 2000US-0195792P.

13-APR-2000; 2000US-0197081P.

14-APR-2000; 2000US-0197087P.

14-APR-2000; 2000US-0197525P.

29-MAR-2001; 2001US-00823187.

(CURA-) CURAGEN CORP.

Majumder K, Spaderna SK, Taupier RJ, Padigar M, Burgess CE;
Shimkets RA, Spytek KA, Liu X, Patturajan M, Gusev VY;
WPI; 2001-626379/72.
N-PSDB; ABA01984.

XX New G protein-coupled receptor related polypeptides and polynucleotides
PT for diagnosis, prevention and treatment of metabolic, neurodegenerative,
PT retinal, immune, hematopoietic disorders, diabetes, obesity and
PT infections.
XX

Claim 1; Page 43; 194pp; English.

XX The present invention provides the protein and coding sequences of novel
CC human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4,
CC NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can
CC be used in the treatment of NOVX related diseases, including cancer,
CC metabolic, neurodegenerative, immune, haematopoietic, developmental,
CC retinal, feeding, neurological and psychotic diseases and disorders and
CC infections. The present sequence is the NOV5a protein, the gene for which
CC is found on chromosome 2. The NOV5a protein shares homology with
CC serotonin receptors
XX

SQ Sequence 370 AA;

Alignment Scores:
Pred. No.: 2,23e-136 Length: 370
Score: 1860.50 Matches: 370
Percent Similarity: 97.63% Conservative: 0
Best Local Similarity: 97.63% Mismatches: 0
Query Match: 86.37% Indels: 9
DB: 4 Gaps: 3

US-09-976-782-15 (1-1152) x AAM47211 (1-370)

Qy 5 ATGGAGGCGCGTAGCCTTTTACGTGCCACCGCGCGTTCCTTGGACCCGAG 64
Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
Qy 65 ACCACAGCGGACCGCGGACCCACCGAGCCGAGGAGTACTGGTTCGACCCGCGGC 124
Db 21 ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 38
Qy 125 GCGGTCTCGCGGCGCGGCGCGCTTCTCTCTTCCACGCTCTCTGTGTGTGAGCGCTG 184
Db 39 AlaValLeuProGlyArgGlyProPheSerValPheThrValLeuValValThrLeu 58
Qy 185 CTAGTGTCTGTGATCGCTGCCACTTTCCTGTGTGAACTCTGTTCCGTCACCAATCCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTyrAsnLeuLeuValProValThrIlePro 78
Qy 245 CCGGTCTCGCTTCCACCGCGTGGCGGATACCTTGGTGGCTCGACCGCGCTCTCGGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
Qy 305 GAACTAGTGGCAGCGCTGGCGATGCCACGAGCTGGCGAGTGAGCTGCGACCGGCGCA 364
Db 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
Qy 365 CGTGGCTGTCTGGCGCGGAGCTGTGCCACGTGTGGATCTCTTTCGACCCGAGCGCTGT 424
Db 119 ArgArgLeuLeuGlyArgSerLeuGlyHisValTyrIleSerPheAspAla----- 135
Qy 425 CTGTGCTGCGCGCGCGCTCGGAACTGGGCGGCGCATCCCTCGCGCGCGCGCGCGC 484
Db 136 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 155
Qy 485 ATCACACGCGCAGCTGCGACACACGCTGCGCACCGCGCGCGCTCTGTGTGTATGATC 544
Db 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 175
Qy 545 GCGCTGCGCGCGGCTCGCGCTCATCGCGCTCATCGCGCTCGCGCGCTCTTTTGGCGCGGC 604
Db 176 AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 195
Qy 605 GAGGTGTGCGACCTCGCTCGGCTCGGCGCTCGGAGTGGAGCGCGGAGACCTCTATGCGCC 664
Db 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerThrAlaAla 215

RESULT 3
ID ABA01985
XX ABA01985 standard; cDNA; 1150 BP.
AC ABA01985;
XX 12-FEB-2002 (first entry)
DT
XX Human NOV5b coding sequence.
DE
XX
XX Human; NOV5b; metabolic disorder; neurodegenerative disorder;
KW immune disorder; haematopoietic disorder; developmental disease; cancer;
KW retinal disease; feeding disorder; vaccine; infection; gene therapy;
KW neurological disorder; psychotic disorder; G-protein coupled receptor;
KW cytosolic; antidiabetic; viricide; neuroprotective; nootropic;
KW analgesic; antidepressant; antimigraine; anticonvulsant; neuroleptic;
KW antiasthmatic; anti-allergic; anti-inflammatory; anorectic; antiarthritic;
KW antipsorathic; antiatherosclerotic; antibacterial; fungicide;
KW osteopathic; protozoacide; antiulcer; hypertensive; hypotensive;
KW antifertility; vulnerary; nephrotropic; antilipemic;
KW serotonin receptor; chromosome 2; ss.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT 5'UTR 1..23
FT FT /*tag= a
FT CDS 24..1136
FT FT /*tag= b
FT FT /*product= "NOV5b"
FT sig_peptide 24..95
FT FT /*tag= c
FT mat_peptide 96..1133
FT FT /*tag= d
FT 3'UTR 1137..1150
FT FT /*tag= e
XX
WO200174851-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US010039.
XX
XX 30-MAR-2000; 2000US-0193205P.
XX 30-MAR-2000; 2000US-0193339P.
XX 05-APR-2000; 2000US-0195343P.
XX 06-APR-2000; 2000US-0195005P.
XX 06-APR-2000; 2000US-0195088P.
XX 10-APR-2000; 2000US-0195792P.
XX 11-APR-2000; 2000US-0196556P.
XX 13-APR-2000; 2000US-0197081P.
XX 14-APR-2000; 2000US-0197087P.
XX 14-APR-2000; 2000US-0197525P.
XX 29-MAR-2001; 2001US-00823187.
XX
(CURA-) CURAGEN CORP.
XX
XX Majumder K, Spaderna SK, Taupier RJ, Padigar M, Burgess CE;
PI Shimkets RA, Spyrek KA, Liu X, Fatturajan M, Gusev VI;
XX
XX WPI; 2001-626379/72.
DR P-PSDB; AAM47212.
XX
XX New G protein-coupled receptor related polypeptides and polynucleotides
PT for diagnosis, prevention and treatment of metabolic, neurodegenerative,
PT retinal, immune, hematopoietic disorders, diabetes, obesity and
PT infections.
XX
XX Claim 8; Page 46; 194pp; English.
XX
XX The present invention provides the protein and coding sequences pf novel
CC human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4,

CC NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can
CC be used in the treatment of NOVX related diseases, including cancer,
CC metabolic, neurodegenerative, immune, haematopoietic, developmental,
CC retinal, feeding, neurological and psychotic diseases and disorders and
CC infections. The present sequence is the NOV5b coding sequence, the gene
CC for which is found on chromosome 2. The NOV5b protein shares homology
CC with serotonin receptors
XX
SQ Sequence 1150 BP; 178 A; 388 C; 355 G; 229 T; 0 U; 0 Other;

Query Match 92.4%; Score 1064.8; DB 4; Length 1150;
Best Local Similarity 97.5%; Pred. No. 5.6e-221;
Matches 1123; Conservative 0; Mismatches 2; Indels 27; Gaps 3;

QY 1 CGCCATGAGGCGCTAGCTTTCAGTGGCCACCGCGCGGTGGCCCTTGGCCCTGGGACC 60
Db 20 CGCCATGAGGCGCTAGCTTTCAGTGGCCACCGCGCGGTGGCCCTTGGCCCTGGGACC 79

QY 61 CGAGACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACCTGGTTCGACCCCGAG 120
Db 80 CGAGACCAGCAG-----CGGACCCCAAGCCGAGAGGATACCTGGTTCGACCCCGAG 133

QY 121 CGGCGCGCTCTGCGGCGCGAGGCGCGCTTCTCTGTCTTACGGTCTCTGGTGTGAC 180
Db 134 CGGCGCGCTCTGCGGCGCGAGGCGCGCTTCTCTGTCTTACGGTCTCTGGTGTGAC 193

QY 181 GCTGCTAGTCTCTGATGCTGCTGCTCTTCTCTGTGGAACTCTGCTGCTGCTACCAT 240
Db 194 GCTGCTAGTCTCTGATGCTGCTGCTCTTCTCTGTGGAACTCTGCTGCTGCTACCAT 253

QY 241 CCGCGGCTCGGTGCTTCCACCGCTGCGGATACCTTGGTGGCTTCCGACCGCTCTC 300
Db 254 CCGCGGCTCGGTGCTTCCACCGCTGCGGATACCTTGGTGGCTTCCGACCGCTCTC 313

QY 301 GGACGAACTAGTGGCAGCGCTGCGATGCCACCGAGCTGGCGAGTGTGCGACCGG 360
Db 314 GGACGAACTAGTGGCAGCGCTGCGATGCCACCGAGCTGGCGAGTGTGCGACCGG 373

QY 361 GCGACGTGCGCTGCTGGCGGAGCTGTGCGACGTGTGATCTCTTCCGCGCGAGC 420
Db 374 GCGACGTGCGCTGCTGGCGGAGCTGTGCGACGTGTGATCTCTTCCGCGCGC 428

QY 421 CTGTCTGTGCTGCCCGCGCGCTCGGGAACGTGGGGCCATCGCCCTGGGCGCGACCG 480
Db 429 ----CTGTGCTGCCCGCGCGCTCGGGAACGTGGGGCCATCGCCCTGGGCGCGACCG 484

QY 481 GGCCATCACAGGCACTGCGACACACGTGGCGACCCCGCAGCGCGCTGCTGTCTCAT 540
Db 485 GGCCATCACAGGCACTGCGACACACGTGGCGACCCCGCAGCGCGCTGCTGTCTCAT 544

QY 541 GATCGGCTCGCCCGGCTGCGGCGCTCATCGGCTCGGCGCTGCTGCTGCTGCGCG 600
Db 545 GATCGGCTCGCCCGGCTGCGGCGCTCATCGGCTCGGCGCTGCTGCTGCTGCGCG 604

QY 601 GGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCCGGGAAACCTCTATGC 660
Db 605 GGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCCGGGAAACCTCTATGC 664

QY 661 CGCCTTCTCACCGCGCGCTTCCACTGCGCTGCGCTGCGCTGCGCTGCTGCTACCG 720
Db 665 CGCCTTCTCACCGCGCGCTTCCACTGCGCTGCGCTGCGCTGCGCTGCTGCTACCG 724

QY 721 GAAGATCTACGAGGCGGCAAGTTTGGTTTGGCGCGCGCGGAGAGCTGTGCTGCGCT 780
Db 725 GAAGATCTACGAGGCGGCAAGTTTGGTTTGGCGCGCGCGGAGAGCTGTGCTGCGCT 784

QY 781 GCGCGGCACTGCGAGGTGAGGTCCAAAGGTAAAGACACCTGATGAGGCTGAAGTGT 840
Db 785 GCGCGGCACTGCG-----AAGTAAAGACACCTGATGAGGCTGAAGTGT 832

QY 841 GTTCCAGGCACATGCAAGCAACCGGTGCTCTTCCAGGTGAGCGGGAATCTCTGGCGGA 900
Db 833 GTTCCAGGCACATGCAAGCAACCGGTGCTCTTCCAGGTGAGCGGGAATCTCTGGCGGA 892

serotonin receptor; chromosome 2; ss.

Homo sapiens.

Key Location/Qualifiers
5'UTR 1..23
/*tag= a
CDS 24..1136
/*tag= b
/*product= "NOV5a"
sig_peptide 24..95
/*tag= c
mat_peptide 96..1133
/*tag= d
3'UTR 1137..1150
/*tag= e

WO200174851-A2.

11-OCT-2001.

30-MAR-2001; 2001WO-US010039.

30-MAR-2000; 2000US-0193205P.

30-MAR-2000; 2000US-0193339P.

05-APR-2000; 2000US-0195343P.

06-APR-2000; 2000US-0195005P.

06-APR-2000; 2000US-0195088P.

10-APR-2000; 2000US-0195792P.

11-APR-2000; 2000US-0196556P.

13-APR-2000; 2000US-0197081P.

14-APR-2000; 2000US-0197087P.

14-APR-2000; 2000US-0197525P.

29-MAR-2001; 2001US-00823187.

(CURA-) CURAGEN CORP.

Majumder K, Spaderna SK, Taupier RJ, Padigar M, Burgess CE;

Shimkets RA, Spytek KA, Liu X, Patturajan M, Gusev VI;

WPI; 2001-626379/72.

P-PSDB; AAM47211.

New G protein-coupled receptor related polypeptides and polynucleotides for diagnosis, prevention and treatment of metabolic, neurodegenerative, retinal, immune, hematopoietic disorders, diabetes, obesity and infections.

Claim 8; Page 42; 194pp; English.

The present invention provides the protein and coding sequences pf novel human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4, NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can be used in the treatment of NOVX related diseases, including cancer, metabolic, neurodegenerative, immune, haematopoietic, developmental, retinal, feeding, neurological and psychotic diseases and disorders and infections. The present sequence is the NOV5a coding sequence, the gene for which is found on chromosome 2. The NOV5a protein shares homology with serotonin receptors

Sequence 1150 BP; 177 A; 388 C; 355 G; 230 T; 0 U; 0 Other;
Query Match 92.7%; Score 1068; DB 4; Length 1150;
Best Local Similarity 97.7%; Pred. No. 1.1e-221;
Matches 1125; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

1 CGCCATGGAGCGCGTAGCCTTTTCAGTGGCCACCGCGCGGTGGCCCTTGCCCTGGGACC 60
20 CGCCATGGAGCGCGTAGCCTTTTCAGTGGCCACCGCGCGGTGGCCCTTGCCCTGGGACC 79
61 CGAGACGAGCGGACCCCGGACCCCAAGCCCGGAGGGATACCTGGTTCGACCCCGGAG 120
80 CGAGACGAGCAG-----CGGGACCCCAAGCCCGGAGGGATACCTGGTTCGACCCCGGAG 133

QY 121 CGGCGCGCTCTCTGCGCGGCGCGAGGCGCGCCCTTCTCTGTCCTTACGGTCTCTGGTGGTAC 180
DB 134 CGGCGCGCTCTGCGCGGCGCGAGGCGCGCCCTTCTCTGTCCTTACGGTCTCTGGTGGTAC 193
QY 181 GCTGCTAGTGTCTGTATGCTGCTGCACCTTTCTCTGTGGAACTGCTGCTGGTCCGCTCACCAT 240
DB 194 GCTGCTAGTGTCTGTATGCTGCTGCACCTTTCTCTGTGGAACTGCTGCTGGTTCGCTCACCAT 253
QY 241 CCGCGGGTCCGTCCTTCCACCGGCTGCCGATAACTTGGTGGCTCGACGGCGCTCTC 300
DB 254 CCGCGGGTCCGTCCTTCCACCGGCTGCCGATAACTTGGTGGCTCGACGGCGCTCTC 313
QY 301 GGACGAACCTAGTGGCAGCGCTGGCGATGCCACCGAGCCTGGCGAGTGTCTCGACCGG 360
DB 314 GGACGAACCTAGTGGCAGCGCTGGCGATGCCACCGAGCCTGGCGAGTGTCTCGACCGG 373
QY 361 GCGAGCTCGGCTGTGGGCGGAGCCTGTGCCACCTGTGGTCTCTCTCGAGCGCGGAGC 420
DB 374 GCGAGCTCGGCTGTGGGCGGAGCCTGTGCCACCTGTGGTCTCTCTCGAGCGCGGAGC 428
QY 421 CTGCTGTGCTGCGCGCGGCTCGGAACTGGCGGCATCGCCCTGGGCGCGGAGCG 480
DB 429 -----CTGTGCTGCCCGCGCGCTCGGAACTGGCGGCATCGCCCTGGGCGCGGAGCG 484
QY 481 GGCCATCACAGCGCACCTTGCGACACACGCTGCGCACCCCGACCGCGCTCGTGTCTCAT 540
DB 485 GGCCATCACAGCGCACCTTGCGACACACGCTGCGCACCCCGACCGCGCTCGTGTCTCAT 544
QY 541 GATCGGCTGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 600
DB 545 GATCGGCTGCGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 604
QY 601 GGGCGAGGTGTGCGAGCGCTGCGGCTCCAGGCTGCGAGGTGAGCGGGAACCTCTCTATGC 660
DB 605 GGGCGAGGTGTGCGAGCGCTGCGGCTCCAGGCTGCGAGGTGAGCGGGAACCTCTCTATGC 664
QY 661 GCGCTTCTCCACCGCGCGCTTCCACCTGCGGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 720
DB 665 GCGCTTCTCCACCGCGCGCTTCCACCTGCGGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGG 724
QY 721 GAAGATCTACGAGGGGCGCAAGTTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 780
DB 725 GAAGATCTACGAGGGGCGCAAGTTTTCGTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGGT 784
QY 781 GCGCGCACCATGCAAGTGAAGTCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGGT 840
DB 785 GCGCGCACCATGCAAGTGAAGTCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGGT 832
QY 841 GTTCAGGCGACATTGCAAGGCAACCGTGTCTTCCAGGTGAGCGGGAATCTCTGGCGGGA 900
DB 833 GTTCAGGCGACATTGCAAGGCAACCGTGTCTTCCAGGTGAGCGGGAATCTCTGGCGGGA 892
QY 901 GCAGAGGAGAGCGGAGGAGCAGCAGTATGTTGGGAATCTGATGGGCTGTTTGTGCTGTG 960
DB 893 GCAGAGGAGAGCGGAGGAGCAGCAGTATGTTGGGAATCTGATGGGCTGTTTGTGCTGTG 952
QY 961 CTGGATCCCTTCTCTTCCAGGAACTCATCAGCCCACTCTGTGCTGAGCGCTGCGCC 1020
DB 953 CTGGATCCCTTCTCTTCCAGGAACTCATCAGCCCACTCTGTGCTGAGCGCTGCGCC 1012
QY 1021 CATCTGGAAGAGCATATTTCTGTGGCTTGGCTACTCAATTTCTTCAACCCCTGAT 1080
DB 1013 CATCTGGAAGAGCATATTTCTGTGGCTTGGCTACTCAATTTCTTCAACCCCTGAT 1072
QY 1081 TTACACAGCTTTTAAACAAGAACTTAAACAATGCTTCAAGAGCTCTTTTACTAAGCAGAG 1140
DB 1073 TTACACAGCTTTTAAACAAGAACTTAAACAATGCTTCAAGAGCTCTTTTACTAAGCAGAG 1132
QY 1141 ATGAACACAGGG 1152
DB 1133 ATGAACACAGGG 1144

Db	127	ATTTCTACCTTGTGGGCTTCTGGTGGGGCGACGTTTGGCTGGAACTCTCTGGTGTG	186
Qy	233	GTACCAATCCCGGGGTCGTTCCTCCACCGCTGCCGATAACTTGGTGGCTCGACG	292
Db	187	GCACCAATCCTCGTGACACCTTCCACCGCTGCCCAACCTGGTGGCATCCATG	246
Qy	293	GCGGTCTCGACGAACTAGTGGACGCTGGCGATGCCACGAGCTGGCGAGTAGCTG	352
Db	247	GCGGTCTCGGATGTCCTGGTGGCGCGCTGGTCAATGCCGTGGAGCTGGTGCACGAGCTG	306
Qy	353	TCGACCGGGCGACGTCGGCTGCTGGGCGGGAGCCTGTGCCACGTGTGGATCTCCTTCGAC	412
Db	307	T---CCGGGCGCGCTGGCAGTAGTTCGGAGGCTGTGCCAGCTTTGGATCGCGTGGAC	363
Qy	413	GCGGAGCCTGTCTGTGTCGCCCGCGCGCTCTGGGAACGTGGCGGCATGCGCCTGGGC	472
Db	364	G-----TGTCTTCGACGCGCGAGCATCTGGAACGCTGACGGGCATAGCACTGGAC	414
Qy	473	CGGACGGGGCCATCACACGSCACCTGCAGCACACGCTGGSCACCGCCAGCGCGCTCG	532
Db	415	CGTACTGTGTCATCAGCGGCCATATGGAATACAGCTCCGCACCCGGAAGTGGTCTCC	474
Qy	533	TTGCTCATGATTCGCGCTTGCOCGGGTGCCGTCGCGCTCATTCGCCCTCGCGCGCTGCTC	592
Db	475	AAGTCAATGATCGCGCTCACCTGGGCACTGNCACCTGTCACTCTCTGGCCCGCTGCTT	534
Qy	593	TTTGGCGGGCGAGGTGGGACGCTCGGCTCCAGCGCTGCCAGCTGAGCGCGGAACCC	652
Db	535	TTTGGCTGGGGAGAGCGTACTCTGAGGGCAGCGAGGAGTGCCAGGTAAAGCGCGAGCCT	594
Qy	653	TCCTATGCCGCCCTTCTCCACCGCGCGCCCTTCACCTGCCGCTTGGCTGGTGCCTGTT	712
Db	595	TCCTACGCGGTCTCTCCA CGTAGGCGCCTTCTACCTGCCGCTCTGTGTGTGCTCTTC	654
Qy	713	GTCTACCGGAAGATCTACGAGGGCGGCAAGTTTCGTTTCGGCGCGCGCGGAGAGCTGTG	772
Db	655	GTGTACTGGGAAGATCTACAAGGCTACCAAGTTCCGCGTGGGCTCCAGGAAGACCAATAGC	714
Qy	773	CTCGCGTTGCGGCGCACCATGACGCTGAGGTCCAAGGTAAAGGAGACACCTGATGAGCT	832
Db	715	GT-----CTCACCATATCGAAGCTGTGGAGGTGAAGGACTCTGCCCAACAGCCC	765
Qy	833	GAAGTGGTGTTCACGSCACATTGCAAAAGCAACGGTGTCTTCCAGGTGAGCGGGGATCC	892
Db	766	CAGATGGTGTTCAGGTCC---GCCACGCCACCGTCACTTCCACGACGAAGGGGACACG	822
Qy	893	TGCGGGGACGAGAAGGAGCGGACAGCCATGATGTGGGAATCTGATTTGGCTGTTT	952
Db	823	TGTGCGGAGCAGAAGGAGCAGCGCGCCCTCATGTGGGCACTCCATTTGGCGTGTTC	882
Qy	953	GTGCTGTGTGGATCCCTTCTTCTTCCAGCGAACTCATACGCCACCTCTGTGCTGCGACG	1012
Db	883	GTGCTGTGTGATCCCTTCTTCTTCCACCGAGCTCATAGTCCCTCTGCTCTGTGAC	942
Qy	1013	CTGCGCCCATCTGGAAGAAGCATTTCTGTGGCTTGGCTACTGCCAATCTTTCTTCAAC	1072
Db	943	ATCCCCGCCATCTGGAAGAAGCATCTCTGTGGCTTGGCTACTCCAACTCTCTTTTAAAC	1002
Qy	1073	CCCTGATTTACACAGCTTTTAA CAAGAACTTACAACAATTCCTTCAAGAGCTCTTTACT	1132
Db	1003	CCCTGATCTACGGGCTTTCACAGAACTACACAGCGCTTCAAGAACTCTCTTTCT	1062
Qy	1133	AAGCA 1137	
Db	1063	AGGCA 1067	

APPLICANT: Amlaiky, No. 5807691rdine
 APPLICANT: Boschert, Ursula
 APPLICANT: Hen, Rene
 APPLICANT: Plassat, Jean-Luc
 TITLE OF INVENTION: Polypeptides Having Serotonin Receptor
 TITLE OF INVENTION: Activity (5H75A), Nucleic Acids Coding for These
 TITLE OF INVENTION: Polypeptides and Uses Thereof
 NUMBER OF SEQUENCES: 13
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Rhone-Poulenc Rorer Inc.
 STREET: 500 Arcola Road, 3043
 CITY: Collegeville
 STATE: PA
 COUNTRY: USA
 ZIP: 19426-0107
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/356.405
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION NUMBER: WO PCT/FR93/00650
 FILING DATE: 29-JUN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: FR 92/08081
 FILING DATE: 01-JUL-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Smith, Julie K
 REGISTRATION NUMBER: 38,619
 REFERENCE/DOCKET NUMBER: EX92004-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610)454-3839
 TELEFAX: (610)454-3808
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1686 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 509..1582
 US-08-356-405-1

Query Match	36.9%;	Score 424.6;	DB 1;	Length 1686;
Best Local Similarity	67.3%;	Pred. No. 4.1e-81;		
Matches 684;	Conservative	0;	Mismatches 309;	Indels 24; Gaps 5
QY	128	GTCTGCGCGGCCGAGGCGCCCTTCTGTCTTACGGTCTCTGGTGTGAGCTGTCTA	187	
Db	590	GTCTGCGCGCCTAGTCGGCCCTTTCTCTCACTTTCGAGTGTAGTCTGACTTTTGTG	649	
QY	188	GTGTGCTGTATCGTGCACATTTCTGTGAACTGCTGTTCCGGTCACTACCCGCG	247	
Db	650	GCGTTTCTAGCTGCGGCCACATTCATTGGAACCTGCTGTGGCTACATCTCAAG	709	
QY	248	GTCCGTGCTCTCCACGCGTGGCCGATAACTTGGTGGCTTCGACGGCGTCTCGAGCAA	307	
Db	710	GTACGCACCTTCCACCGAGTACACACAACTGGTAGCTTCCATGGCCATCTCGATGTG	769	
QY	308	CTAGTGGCAGCGCTGGGATGCCACCGAGCCTGGCGAGTGAGCTGTGACCGGCGACGT	367	
Db	770	CTAGTGGCTGTGTGTTATGCCACTGAGCTGGTGCATCATGAGCTGT--CTGGGGCGCGC	826	
QY	368	CGCGTCTGGGCGGAGCGCTGTGCCAGCTGTGGATCTCTTGACGCGCGGAGCGTCTGTG	427	
Db	827	TGGCAGCTGGGCGGCGCTGTATGCCAGCTGTGATGCATGTGACG-----TGCTC	877	

RESULT 3
US-08-356-405-1
; Sequence 1, Application US/08356405
; Patent No. 5807691
; GENERAL INFORMATION:

Qy	428	TGCTGCCCCCGGCTCTGGGAACGTGGCGGCATCGCCCTGGGCGCGAGCGGGCAATC	487
Db	878	TGCTGTACTGCGAGCATCTGGAATGTACACAGCAATAGCACTGGACCGCTACTGCTCAATC	937
Qy	488	ACACGGCACCTGGAGCACAGCTGCGCACCGGAGCGCGCTCTGTTGCTCATGATCGG	547
Db	938	ACGGCACCTGGAGTACACACTCGGTACCGCAAGCGTGTCTCCAATGTGATGATCTGTG	997
Qy	548	CTCGCCCGGGTGGCGTGGCGCTCATCGCCCTCGCGCGCTGTCTTTGGCCGGGGCGAG	607
Db	998	CTACCTGGGCACCTCTCCACTGTTCATCTCTGTGCTCCACTGTATTTGGCTGGGAGAG	1057
Qy	608	GTGTGCGACGCTCGGCTTCGAGCGTGCAGGTGAGCGGAAACCTCTCATGCGCGCTTC	667
Db	1058	ACTTATTCTGAGCCCGAGTGAAGAAATGCCAGTCAGTGCGAGCGCTTCTCTACACCGTGTTC	1117
Qy	668	TCCACCCGCGCGCTTCCACCTGCCCTCTGGCGTGTGCGTGTGCTTGTCTACCGGAAGATC	727
Db	1118	TCCACCGTGGTGCTTCTACCTGCCGTGTGGCTGTGTCTTTGTGTACTGGAAAAATT	1177
Qy	728	TACAGAGGGCCAGATTTCGTTTGGCGCGCGCGGAGAGCTGTGCTGCCGTTGCCGGCC	787
Db	1178	TACAGGGCGGGAATTCGCGATGGGTCCAG-----GAAGACTAACAGCGTCTCCCCCG	1232
Qy	788	ACCATGCAAGTGAGGTCACAGETAAAGGAAGCACCTGATGAGGTGAAGTGGTGTTCACG	847
Db	1233	TACCGGAAGCTGTGG-----AGGTGAAGATGCTACACAACATCCCAGATGGTGTTCACG	1288
Qy	848	GCACTTTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTGGCGGGAGCAGAAG	907
Db	1289	GCCC---GCCATGCCACCGTCACCTTCCAGACAAGGGGATACGTGGAGGGAGCAGAAG	1345
Qy	908	GAGAGGCGAGCAGCATCATGCTGGGAATTCGTATTGGCGTGTGTGCTGTGCTGGATC	967
Db	1346	GAGCAAGGGCAGCCCTCATGTGGGCATCTCTATCGAGGTGTGTGCTCTGCTGTCTC	1405
Qy	968	CCCTTCTTCTGACGGAACATCATAGGCCACTCTGTGCTGTGACGCTTGGCCCCCAATCGG	1027
Db	1406	CCTTCTCTCGTACAGAGCTCATCAGTCCCTGTGTTTCTTGGGACGTCCCTGCCATCTGG	1465
Qy	1028	AAAAGCATATTTCTGTGGCTTGGCTACTCCAATTTCTTCTTCAACCCCTGATTACACA	1087
Db	1466	AAGAGCATTTCTCTGTGGTATTTCTAAATTTCTTCTTCAACCCCATCTCTACACA	1525
Qy	1088	GCCTTTTAAAGAACTACAACAATGCCTTCAAGAGCCTCTTTTACTAAGCAGAGATGA	1144
Db	1526	GCATTTCAACAGAGCTACAGCAGTGTCTTCAAGTCTTCTTCTCTCAAGCAACAATGA	1582

RESULT 4

US-08-031-538-3
US-08-031-538-3
Sequence 3, Application US/08031538
Patent No. 5968817
GENERAL INFORMATION:
APPLICANT: Sutcliffe, J Gregor
APPLICANT: Brlander, Mark G
APPLICANT: Lovenberg, Timothy W
TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
NUMBER OF SEQUENCES: 73
CORRESPONDENCE ADDRESS:
ADDRESS: The Scripps Research Institute, Office of
ADDRESS: Patent Counsel
STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25


```

; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1077:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g1162923
;
US-09-016-434-1077

Query Match 8.3%; Score 95.2; DB 4; Length 1984;
Best Local Similarity 47.1%; Pred. No. 1.9e-11;
Matches 489; Conservative 0; Mismatches 513; Indels 36; Gaps 5;

QY 114 CCCGAGCGGCGCGTCTCGCGGGCCGAGGGCGCCCTTCTCTGCTTTCACGGTCCGTGG 173
Db |||||
QY 493 CCAATAGCACCCTGGGGGGCAGGGCGCGCTCGGCCCGGGGGCAGCGCTGGG 552
Db |||||
QY 174 TGGTGACCTGTAGTGTGCTGATCGCTGCACCTTTCTGTGGAACCTGTGGTTCCGG 233
Db |||||
QY 553 TGGCGGCGCGCTGTGCGTGTGTCATCGCGTGCAGGGGGCGCCAACTCGCTGTGATCG 612
Db |||||
QY 234 TCACCATCCCGGGGTCGCGCTTCCACCGGTCGCGCAAACTTGTGTGGCTCGACGG 293
Db |||||
QY 613 CGCTCATCTGCACTCAGCGCGCGCTGCGCAACACAGCTCCAACTTCTCTGCTGTGCTCT 672
Db |||||
QY 294 CCGTCTCGAGCAACTAGTGTGCGAGCGCTGGGGATGCCACCGAGCTGCGGAGTGAGTGT 353
Db |||||
QY 673 TCAGCTGCACTGATGTGGGGCTGGTGGTGTATGCCCGCGCCATGCTGAACGCGCTGT 732
Db |||||
QY 354 CGACCGGCGAGCTGCGCTGTGGCGCGGAGCTGTGCAAGTGTGATCTCTTTGACG 413
Db |||||
QY 733 ACG-----GGCGTGGTGTGCGCGCGGCTCTGCTGTCTGGACCGCTTCGACG 786
Db |||||
QY 414 CCGGAGCCTGTCTGTGCTGCCCGCGCGCTCGGGAAAGTGCGGCCCATCGCCCTGGGCC 473
Db |||||
QY 787 -----TGATGTGTGCGAGCGCTCCATCTCAAGCTCTGCGCTCATCAGCCTGGACC 837
Db |||||
QY 474 GCGACGGGGCATCACACGGCACCTGCGACGACGCTGCGACCGCGAGCGCGCTCGT 533
Db |||||
QY 838 GCTACTGCTCATCTCTGCGCGCTGCGCTACAGCTGGGATGAGCCCTCGCGTGGCC 897
Db |||||
QY 534 TGCTCATGATCGCGCTCGCGCGGTGCGGTGCGGCTCATCGCCCTCGCGCGCTGCTCT 593
Db |||||
QY 898 TGGCCCTAGTCTCGGGCGCTGAGGCTCGCGCTCTCGCTCTCTTCTGCGCCCTGTGCG 957
Db |||||
QY 594 TTGGCGGGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGTGTGAGCGGGAACTT 653
Db |||||
QY 958 TGGGCTGCGAGAGTGGGGCCACGACGCGC-----CACCGTCCCTCGG 1000
Db |||||
QY 654 CCTATGCGCGCTTCTCCACCGCGCGCTTCCACTGCGGTGTGGCGTGGTGGCGTTTG 713
Db |||||
QY 1001 CAGTGGCGCGCTGTGGCGAGCTGCTTTTGTCTTGTGGGTGCGGCGCTCACCTTCTT 1060
Db |||||
QY 714 TCTACCGGAAGATCTACAGGCGGCGCAAGTTTCGTTTCGCGCGCGCGGAGAGTGTGCG 773
Db |||||
QY 1061 CTTGCCCTCGGGTGCCATATGCTTACCTACTGCGAGGATCTCTGTAGTGTGCCCGC-AAAG 1119
Db |||||
QY 774 TGCCGTTGCCCGCACCATGCGAGGTGAGTCCAAAGTAAGAACACTGATGAGGCTG 833
Db |||||
QY 1120 AGGCGGTGAGGTGGCTTCCCTCACACCGGCGATGGCGAGTCAGGCTCGGAGAGCTGCG 1179
Db |||||
```

```

QY 834 AAGTGTGTTCACGGCACATTTCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCT 893
Db |||||
QY 1180 AGGTGCCAGGACCCCAAGCCAGGGGTGGAGTCTGCTGACAGAGCGGTCTAGCCACGA 1239
Db |||||
QY 894 GCGGGAGCAGAGGAGGAGGAGCCATGATGTTGGGGAATTCGATTTGGCGTGTGTTG 953
Db |||||
QY 1240 AGCAGCAGGAGGCGCTGAAAGCCAGCCCTGACGCTGGGCATCTGCTGGGCATGTTCT 1299
Db |||||
QY 954 TGCTGTGTGATCGCTTCTTCTGACGGAACATCATCAGCCCACTCTGTGCTGCGAGCC 1013
Db |||||
QY 1300 TTGTGACCTGGTGGCCCTTCTTGTGGCCAACTAGTCCAGGCGGTGTGCGACTGCTCT 1359
Db |||||
QY 1014 TGCCCCCATCTGGAAGACATATTTCTGTGGCTTGGCTACTCCAAATTTCTTTCAACC 1073
Db |||||
QY 1360 CCCAGGCGCTTCGA---TGTCTCATATGGCTGGGTACTGTAAACAGCACCATGAACC 1416
Db |||||
QY 1074 CCCTGATTACAGCTTTTAAACAAGAACTACAAATGCTTCAAGAGCCTCTTTACTA 1133
Db |||||
QY 1417 CCATCATCTACCCACTCTTTCATGCGGGACTTCAAGCGGCGCTGGGCAAGTTCCTGCCAT 1476
Db |||||
QY 1134 AGCAGATGAACACAGG 1151
Db |||||
QY 1477 GTCCACGCTGTCCCGGG 1494
Db |||||

RESULT 7
US-09-016-434-1197
; Sequence 1197, Application US/09016434
; Patent No. 6500938
; GENERAL INFORMATION:
; APPLICANT: Janice Au-Young
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
; TITLE OF INVENTION: PATHWAY GENE EXPRESSION
; NUMBER OF SEQUENCES: 1490
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING DATE: HEREWITH
; APPLICATION NUMBER: US/09/016,434
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0002 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1197:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1756 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g181431
;
US-09-016-434-1197
```


Db 710 CAAGAACCTAGAGCGAGGAGTCTAGAGGAGATGTCCAACTCCAGAGGAGCTGACCCTG-- 767
QY 811 AAAGGAAGACCTGATGAGGCTGAAGTGGTGTTCACGGGCAATTCGAAAGCAAGGTTGTC 870
Db 768 --AGGATCCATTCCAGAACTTTTCACGAGGACACCTTTAGCAGTACCAAGGCCAAGGGCC 825
QY 871 CTTCCAGGTAGCGGGGACTCCTGGCG-----GGAGCAAGAGGAGGCGAGCAG 920
Db 826 ACAACCCAGGAGTTCATAGCTGTCAAACTTTTAAAGTTCTCCAGGGAAGAAAGACAG 885
QY 921 CCATGATGGTGGGAATTCGATTTGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 980
Db 886 CTAAGAGCTTGGGCAATGTGGTGGTATGTTCATCTTGTGCTGCTACCTCTTCTATCG 945
QY 981 CGGAATCATACAGCCCACTCTGTGCGCTGCGAGCTGCCCCC---CATCTGGAAGAGCATAT 1037
Db 946 CTCCTACCGCTTGGCTCTCTTCTCTCCACCTGAAGCCCGCCGAGCGCGTGTCAAGGTGG 1005
QY 1038 TTCTGTGGCTTGGCTACTCCAAATCTTTCTCAACCCCTGATTTACAGCTTTTAAACA 1097
Db 1006 TGTCTGTGGCTGGGCTACTTCAACAGCTGCGCTCAACCCCAATCATCTCCATGCTCCAGCA 1065
QY 1098 AGAACTACAAATGCTTCAAGAGCCTCTTT 1129
Db 1066 AGGAGTTCAAGCGGCTTTCATGGTATCCTT 1097

RESULT 12

US-08-351-473B-7
; Sequence 7, Application US/08351473B
; Patent No. 5656440
; GENERAL INFORMATION:
; APPLICANT: LENZEN, GERLINDA
; APPLICANT: KAPOOR, ARCHANA
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES CODING FOR THE
; TITLE OF INVENTION: BOVINE BETA3-ADRENERGIC RECEPTOR AND THEIR APPLICATIONS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/351.473B
; FILING DATE: 21-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 93 04670
; FILING DATE: 21-APR-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/FR94/00447
; FILING DATE: 21-APR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 6639-001-0X PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1227 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-351-473B-7
Query Match 7.8%; Score 89.8; DB 1; Length 1227;
Best Local Similarity 48.6%; Pred. No. 2.4e-10;
Matches 460; Conservative 0; Mismatches 442; Indels 45; Gaps 6;
QY 165 CGGTCTCTGGTGGTGAAGCTCTAGTGTCTGATCGTGCACCTTCTCTGTGGAACCTGC 224
Db 110 CGGCCCTTAGCCGGGGCCCTCTGCGCTGCGGTGTGGCCACCGTGGGAGGCAACTGTC 169
QY 225 TGGTTCCTGGTCAACATCCCGCGGTCCGTCCTCCACCGCGTGCACATAACTTGGTGG 284
Db 170 TGGTCATCTGGCCATCGCTCGACTCCGAGACTCCAGACCATGACCAACGTGTTCGTGA 229
QY 285 CTTTCGAGCGCGCTCTCGAGCAACTAGTGGCAGGCTGGCGATGCGCACCGAGCCTTGGCGA 344
Db 230 CTTTCGCTGGCGCGAGCGCACTGGTGTATGGCACTCTCTGGTGGTGGCGCC-----GGCGG 283
QY 345 GTGAGCTGTGACCGGGCGAGCTCGGCTGCTGGGCGGAGCCTGTGCGACAGTGTGGATCT 404
Db 284 CCACCTTGGCGCTGACTGGCCACTTGGCCGTTGGCGCCACTGGCTGCGAGCTGTGACCT 343
QY 405 CTTTCGAGCGCGGAGCCTGTCTGTGCTGCCCGCGGCTCGGGAAGCTGGCGGCAATCG 464
Db 344 CGGTGGAGCTG-----CTGTGTGACCGCGCAGCATCGAAACCTGTGTGCGCCCTGG 394
QY 465 CCTTGGCGCGGAGCGGCGCATACACGCGCACTGAGCAACGCTGGGCAACCGCGAGCC 524
Db 395 CCGTGGAGCGCTACTGGCTGTGACCAACCGCTGCGTTAGCGGGCACTGGTCAACCAAGC 454
QY 525 GCGCTCTGTGCTCATGATCGGCTCGCGGCTCGCGGTGCGGCTCATCGCCCTCGGCG 584
Db 455 GCTGCGCGCGGACAGCTGTGCTCTGTGTGGGTGCTGTGCGCGCGGTGTGTTTGGCG 514
QY 585 C---GCTGCTCTTTGGCGGGGAGGTGTGCGACGCTCGGCTCCAGCGCTCGCAGGTG 641
Db 515 CCATCATGAGCGCAGTGGTGGCGGTAGGGGCCGACGCGGAGCGCAGCGCTGCCACTCA 574
QY 642 GCGGGGAACCTCTATTCGCGCT-----TCTCCACCGCGGCG 680
Db 575 ACCCGCTGCTGTGCGCTTGGCTTCCAAACATGCGCTAGCTGCTGTCTCTCTCCGCT 634
QY 681 CTTTCACCTGCGCTTGGGTGGTGGCTTGTGCTACCGGAAGATCTACGAGCGGCA 740
Db 635 CTTTCTACCTTCTCTCTGCTGATGCTCTTCTGCTAGCGCGGCTTCTGTTGGCTA 694
QY 741 AGTTTCGTTTCGCGCGCGCGGAGAGCTGTGCTGCGGCTTGGCGGCGCACCATGAGGTGA 800
Db 695 CGGCCAGCTGCGCTTGTGCGCGGGAGCTGGGCGCTTTCG---CCGAGGAGTCTC 751
QY 801 GGTCCAAAGGTAAAGGAAGCACTGATGAGGTGAAGTGGTGTTCACGGCAATTCGAAG 860
Db 752 GCGCGGCGCGCTGCGGCTCTCTGCGCGCGCGCGCTGGGACGCTGCGCTCGCGCCGAAG 811
QY 861 CAACGCTGTCTTCCAGGTGAGCGGAGCTCTGCGCGGAGCAGAGAGAGAGGAGAGCAG 920
Db 812 GGGTGGCGGCTGCGGCGCGGCGCGGCGCTCTGCGCTCTCGGGAACACCGGGGCC 871
QY 921 CCATGATGGTGGGAATTCGATTTGGCGTGTGTGTGCTGTGATGATGCTTCTTCTCTGA 980
Db 872 TGTGCACTTGGGTCTCATATGGGCACTTCACTCTCTGCTGCTGCTTCTTCTTCTGG 931
QY 981 CGGAACCTCATAGCCCACTCTGTGCGCTGAGCCTGCGGCGGCGGCTGCGGAAGCATTTTC 1040
Db 932 CCAACGCTGCTGCGCGCTCTGGGCGGCGCTCTCTAGTCCCGGCGCGCTTCTCTCTGCC 991
QY 1041 TG---TGGCTTGGCTACTCCAAATCTTCTTCAACCCCTGATTTAC 1084
Db 992 TGAACCTGGCTAGGTTATGCCAAATCTGCTTCAACCCCGCTCATCTAC 1038

Query Match	7.7%; Score 88.8; DB 1; Length 1738;
Best Local Similarity	48.9%; Pred. No. 4.2e-10;
Matches 366; Conservative	0; Mismatches 362; Indels 20; Gaps 4;
QY	394 CGGTGGATCTCTTCGACCGCGAGCCTCTCTGTGTCGCCGCCGCCGCTCGGACGT 453
Db	471 CTTCTGTGACATCTGGCAGCGGTGATGCTCTGTCTGCACACGCTCCATTCTGAGCCT 530
QY	454 GCGGGCCATCGCCCTTGGCGCGCAGCGGGCCATCACACGGCACCTGCAGACACACGCTGCG 513
Db	531 GTGCGCCATCTCCATCGATCGCTACATCGGGTCCGCTACTCTCTCGAGTATCCACGCT 590
QY	514 CACCGGACCGCGCCTCGTTGCTCATGATCGCGCTCGCCGGGTGCGCTCGGGCGCTCAT 573
Db	591 GGTCAACCGGAGAAAGCCATCTTGGCGCTGCTCAGTGTCTGGGTCTTGTTCCACCGCTCAT 650
QY	574 CGCCCTCGCGCGCTGCTCTTTGGCGGGGCGAGGTGTGACGCTCGGCTCCAGCGCTG 633
Db	651 CTCATTCGGGCTCTCTTGGTGGAGGACCGGCACCCACGAT---GACAGGAGTG 707
QY	634 CAGGTTAGCCGGGAACCCCTCTTATGCGCGCTTCTCCACCGCGGGCGCCTTCCACCTGCC 693

```
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS
LOCATION: 124..1683
OTHER INFORMATION:
US-08-228-932-3

Query Match
Best Local Similarity 7.7%; Score 88.8; DB 1; Length 1738;
Matches 366; Conservative 0; Mismatches 362; Indels 20; Gaps 4;

QY 394 CGTGTGGATCTCTTCGAGCGGAGCGCTGTCTGTGCTGCCCCCGGCGCTCGGGAACGT 453
Db 471 CTTCTGTGACATCTCGGCGAGCGTGGATGCTCTGTGCTGCACAGCGTCCATTCTGAGCT 530
QY 454 GCGCGCCATCGCCCTGGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 513
Db 531 GTGCGCCATCTCCATCGATCGCTACATCGGGGTGCGTACTCTCTGAGTATCCACGCT 590
QY 514 CACCGCGAGCGCGCTCGTGTCTCATGATCGCGTGCGCCGGGTGCGCGTCCGCTCAT 573
Db 591 GGTCAACCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 650
QY 574 CCGCGTGGGCGTCTCTTCGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 633
Db 651 CTCATCGGCGCTCTCTTCGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 707
QY 634 CCAGGTAGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 693
Db 708 CGGGTCAACCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 767
QY 694 GCTTGGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 753
Db 768 TCTGCGCGTCACTTAGTCACTAGTCACTAGTCACTAGTCACTAGTCACTAGTCACTAGT 827
QY 754 CCGCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 813
Db 828 GAACCTAGAGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 883
QY 814 GGAAGCACTGATGAGCGTGAAGTGTGTTCACGGCAGGCGGCGGCGGCGGCGGCGGCGGCGG 873
Db 884 GGATCACTTCAAGAACTTTCACGAGGACACCTTAGCAGTACCAAGCGGCGGCGGCGGCGG 943
QY 874 CCAGGTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 923
Db 944 ACCCGAGGAGTTCATAGCTGTCAAACTTTTAAAGTTCCTCAGGGAAGAAAGAGCGTA 1003
QY 924 TGATGTGGGAACTTCTGATTCGCGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 983
Db 1004 AGAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1063
QY 984 AACTCATCAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1040
Db 1064 TACCGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1123
QY 1041 TGTGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1100
Db 1124 TCTGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1183
QY 1101 ACTACAACAAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1128
Db 1184 AGTTCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1211
```

RESULT 15

US-08-468-939-3
; Sequence 3, Application US/08468939
; Patent No. 5714381
; GENERAL INFORMATION:

```
APPLICANT: Jonathan A. Bard et al.
TITLE OF INVENTION: DNA Encoding Human Alpha 1 Adrenergic
NUMBER OF INVENTIONS: 6
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: COOPER & DUNHAM LLP
CITY: New York
STREET: 1185 Avenue of the Americas
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/468,939
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 41337-1B
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0526
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1738 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: N
ANTI-SENSE: N
FEATURE:
NAME/KEY: CDS
LOCATION: 124..1683
OTHER INFORMATION:
US-08-468-939-3
```

```
Query Match
Best Local Similarity 7.7%; Score 88.8; DB 1; Length 1738;
Matches 366; Conservative 0; Mismatches 362; Indels 20; Gaps 4;

QY 394 CGTGTGGATCTCTTCGAGCGGAGCGCTGTCTGTGCTGCCCCCGGCGCTCGGGAACGT 453
Db 471 CTTCTGTGACATCTCGGCGAGCGTGGATGCTCTGTGCTGCACAGCGTCCATTCTGAGCT 530
QY 454 GCGCGCCATCGCCCTGGGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 513
Db 531 GTGCGCCATCTCCATCGATCGCTACATCGGGGTGCGTACTCTCTGCAAGTATCCACGCT 590
QY 514 CACCGCGAGCGCGCTCTGTGCTCATGCTGCGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGG 573
Db 591 GGTCAACCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 650
QY 574 CCGCGTGGGCGTCTCTTCGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 633
Db 651 CTCATCGGCGCTCTCTTCGAGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 707
QY 634 CCAGGTAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 693
Db 708 CGGGTCAACCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 767
QY 694 GCTTGGCGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 753
Db 768 TCTGCGCGTCACTTAGTCACTAGTCACTAGTCACTAGTCACTAGTCACTAGTCACTAGT 827
QY 754 CCGCGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 813
Db 828 GAACCTAGAGCGGAGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 883
```

```

QY 814 GGAAGCACCCTGATGAGGCTGAAGTGTGTTACGGCACATTGCAGACAAAGCAACGGTGTCTTT 873
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 884 GGATCCATTCCAGAACTTTCCAGGACACCTTTAGCAGTACCAAGGCCAAGGGCCACA 943
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 874 CCAGGTGAGCGGGGACTCCTGGCG-----GGAGCAGAAGGAGGCGAGCAGCCCA 923
Db | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 944 ACCCCAGAGTTCCATAGCTGTCAAACTTTTAAAGTTCTCCAGGGAAAAGAACAGCTA 1003
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 924 TGATGTGGGAATCTGATGGCGTGTGCTGCTGGATCCCTTCTTCCTGACGG 983
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1004 AGACGTTGGGCATTTGGTGGTATGTTCACTTGTGCTGGCTACCCCTTCTTCATCGCTC 1063
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 984 AACTCATCAGCCCACTCTGCTGCAGCTGCCCC---CATCTGAAAAGCATATTTC 1040
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1064 TACCGCTTGGCTCTTGTCTCACCCTGAAGCCCCCGACGCCGTGTTCAAGGTGGTGT 1123
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1041 TGTGGCTTGGCTACTCCAATTTCTTCTCAACCCCTGATTTACACAGCTTTTAAACAAGA 1100
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1124 TCTGGCTGGCTACTTCAACAGCTGCTCAACCCCATCATCTACCCATGCTCCAGCAAGG 1183
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1101 ACTACAAATGCTTCAAGAGCTCTT 1128
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1184 AGTTCAAGCGCGTTTCGTGGCATCCT 1211
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Search completed: September 7, 2004, 17:15:10
Job time : 114 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 14:14:05 ; Search time 494 Seconds
(without alignments)
9906.729 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence: 1 cgcacgagcgccgtagcc.....aagcagagatgaacacacagg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq_29Jan04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002s.*

7: Geneseqn2003as.*

8: Geneseqn2003bs.*

9: Geneseqn2003cs.*

10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	100.0	1152	6	ABN86917 Human NOV
2	1068	92.7	1150	4	ABA01984 Human NOV
3	1064.8	92.4	1150	4	ABA01985 Human NOV
4	1064	92.4	1458	6	AAD29681 Human G-p
5	1060	92.0	1119	6	AAD24963 Human G-p
6	1016	88.2	1155	6	ABK71923 Human cDN
7	756.2	65.6	3086	9	ADC86156 Human GPC
8	745	64.7	977	6	AA898053 Human DNA
9	732.8	63.6	2036	2	AAQ70264 Murine se
10	732.8	63.6	2061	7	AAQ48747 Mouse 5-H
11	729.8	63.4	2226	2	AAQ72270 Rat MR22
12	729.8	63.4	2240	9	ADB52419 Primary r
13	721.2	62.6	1090	6	AA898145 Human DNA
14	497.6	43.2	1074	6	ABK50434 Human 5-h
15	497.6	43.2	1074	7	ABZ42828 Human 5-H
16	497.6	43.2	1159	9	AA58496 Human 5-h
17	496	43.1	2700	6	ABK34848 Human cDN
18	492.8	42.8	1074	5	AB197963 Non-endog
19	484.4	42.0	1073	2	AAQ56309 Human SHI
20	435.2	37.8	1112	6	ABQ43624 Oligonuc1
21	435.2	37.8	1112	6	ABQ43625 Oligonuc1
22	424.6	36.9	1686	6	ABK13748 Mouse 5-H
23	421.4	36.6	1686	2	AAQ56308 Murine 5H

24	416	36.1	2155	2	AAQ72269 Rat RECI7
25	390.2	33.9	1112	6	ABQ43623 Oligonuc1
26	390.2	33.9	1112	6	ABQ43622 Oligonuc1
27	378.2	32.8	636	4	AAH42175 Nucleotid
28	334.4	29.0	336	2	AAQ70265 Human ser
29	327.8	28.5	419	6	AA898116 Human DNA
30	320	27.8	796	3	AA871992 Single nu
31	319.6	27.7	796	3	AA871995 Single nu
32	319.6	27.7	796	3	AA871989 Single nu
33	318.4	27.6	2543	6	ABK50432 Human 5-h
34	311.4	27.0	354	7	ABZ81690 Human LP3
35	297	25.8	297	4	AAH42174 Nucleotid
36	259.4	22.5	273	7	ABZ81688 Polynucle
37	211.4	18.4	225	7	ABZ81689 Polynucle
38	203.6	17.7	1929	6	ABK50433 Human 5-h
39	142.4	12.4	200	7	AA848749 Targettin
40	125.4	10.9	1140	2	AA88392 Corn barn
41	106	9.2	2826	7	ABZ42625 Human alp
42	104.4	9.1	1382	2	AAQ64890 Human der
43	104.4	9.1	1382	6	ABZ35643 Human gen
44	104.4	9.1	1382	7	ACA56658 Human sig
45	102.8	8.9	1371	4	AA199933 Human alp

ALIGNMENTS

RESULT 1
ABN86917

ID ABN86917 standard; cDNA; 1152 BP.

AC ABN86917;

DT 29-JUL-2002 (first entry)

DE Human NOV5 encoding cDNA sequence SEQ ID NO:15.

Human; NOV5; cytostatic; antiarteriosclerotic; cardiovascular; lymphoma;
antidiabetic; immunosuppressive; neuroprotective; gene therapy; cancer;
cardiomyopathy; atherosclerosis; cell signal processing; diabetes; AIDS;
metabolic pathway modulation; neoplastic; neurological disorder; asthma;
adenocarcinoma; prostate cancer; uterus cancer; immune response;
Crohn's disease; multiple sclerosis; Graft versus host disease;
chromosome 2; gene; ss.
Homo sapiens.

Key Location/Qualifiers
CDS 5..1144
FT /*tag= a
FT /product= "NOV5"
FT /note= "Serotonin receptor-like protein"

WO200230974-A2.

18-APR-2002.

12-OCT-2001; 2001WO-US031922.

12-OCT-2000; 2000US-0240113P.

16-OCT-2000; 2000US-0240625P.

16-OCT-2000; 2000US-0240637P.

16-OCT-2000; 2000US-0240648P.

16-OCT-2000; 2000US-0240662P.

16-OCT-2000; 2000US-0240703P.

16-OCT-2000; 2000US-0240732P.

18-JAN-2001; 2001US-0262455P.

(CURA-) CURAGEN CORP.

(MILL/) MILLET I.

PI Grosse WM, Alsobrook JP, Lepley DM, Burgess CE, Mishra V;
 PI Kekuda R, Li L, Padigar M, Shimkets RA, Zerhusen BD, Spytsek KA;
 PI Edinger S, Gerlach V, Macdougall J, Stone D, Gunther E, Ellerman K;
 XX P-PSDB; ABB78809.
 DR WPI; 2002-444172/47.
 DR
 XX
 DR
 XX
 PT New NOVX polypeptides and polynucleotides, useful for treating or
 PT preventing a NOVX-associated disorder or a pathological state in a
 PT subject, particularly a human, e.g. cardiomyopathy, atherosclerosis,
 PT cancer or diabetes.
 XX
 PS
 PS Claim 9; Page 46; 227pp; English.
 XX
 CC The present invention describes novel human proteins designated NOVX
 CC (where X is 1, 2a, 2b, 3, 4, 5, 6a, 6b, 7, 8, or 9). NOV1 is a
 CC tyrosine-protein kinase 6-like protein; NOV2a-d are keratin 4-like
 CC proteins; NOV3 is a collagen-like protein; NOV4 is a cystatin B-like
 CC protein; NOV5 is a serotonin receptor-like protein; NOV6a and NOV6sv are
 CC cold inducible glycoprotein 30-like proteins; NOV7 is a matrilin-2-like
 CC protein; NOV8 is a leukocyte surface antigen (CD53)-like protein; and
 CC NOV9 is a tyrosine kinase-like protein. NOVX sequences have cytostatic,
 CC antiatherosclerotic, cardiovascular, antidiabetic, immunosuppressive and
 CC neuroprotective activities, and can be used in gene therapy. The NOVX
 CC sequences can be used in therapeutics, particularly for treating,
 CC preventing or alleviating a NOVX-associated disorder or a pathological
 CC state in a subject, particularly a human. These disorders include
 CC cardiomyopathy, atherosclerosis, a disorder related to cell signal
 CC processing and metabolic pathway modulation or diabetes. The NOVX
 CC sequences are also useful for determining the presence of or
 CC predisposition to a disease associated with altered levels of NOVX
 CC polypeptide or nucleic acid, particularly cancer. The NOVX sequences are
 CC especially useful in therapeutic or prophylactic applications for
 CC neoplastic or neurological disorders, and in the treatment of
 CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
 CC response, AIDS, asthma, Crohn's disease, multiple sclerosis or Graft
 CC versus host disease. The present sequence encodes the human NOV5 protein
 CC from the present invention. NOV5 is located to chromosome 2
 XX
 XX Sequence 1152 BP; 176 A; 388 C; 359 G; 229 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 1152; DB 6; Length 1152;
 Best Local Similarity 100.0%; Pred. No. 7, 4e-240;
 Matches 1152; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGGCATGAGGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTGCCCTTGCCCTGGGACC 60
 DB 1 CGGCATGAGGCGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTGCCCTTGCCCTGGGACC 60
 QY 61 CGAGACACGAGCGGACCGGGACCCCAAGCCGAGAGGGATCTCGGTTTCGACCCCGAG 120
 DB 61 CGAGACACGAGCGGACCGGGACCCCAAGCCGAGAGGGATCTCGGTTTCGACCCCGAG 120
 QY 121 CGGCGCGCTCTGCGCGGCGGAGGCGCGCCCTTCTGTCTTTCACCGTCTGTGTGTGAC 180
 DB 121 CGGCGCGCTCTGCGCGGCGGAGGCGCGCCCTTCTGTCTTTCACCGTCTGTGTGTGAC 180
 QY 181 GCTGTAGTGTCTGATGCTGCGACTTTCCTGTGTGGAACCTGCTCGGTTCGGTCAACAT 240
 DB 181 GCTGTAGTGTCTGATGCTGCGACTTTCCTGTGTGGAACCTGCTCGGTTCGGTCAACAT 240
 QY 241 CCGCGGCGTTCGTCCTTCCACCGGTGCGGATAACTTGGTGGCTTCGACGCGCGTCTC 300
 DB 241 CCGCGGCGTTCGTCCTTCCACCGGTGCGGATAACTTGGTGGCTTCGACGCGCGTCTC 300
 QY 301 GGAACGAACCTAGTGGAGCGCTGGCGATGCGACCGAGCTGGCGAGTGAGTGTGACCGG 360
 DB 301 GGAACGAACCTAGTGGAGCGCTGGCGATGCGACCGAGCTGGCGAGTGAGTGTGACCGG 360
 QY 361 GCGACCTCGGCTGCTGGCGCGGAGCGCTGTGCGACGTGTGATCTCTTCGACGCCGAGC 420
 DB 361 GCGACCTCGGCTGCTGGCGCGGAGCGCTGTGCGACGTGTGATCTCTTCGACGCCGAGC 420

QY 421 CTGTCTGTGCTGCTGCGCGCGCGCTCGGGAACCTGCGGCAATCGGCCCTGGGCGGACCG 480
 DB 421 CTGTCTGTGCTGCTGCGCGCGCGCTCGGGAACCTGCGGCAATCGGCCCTGGGCGGACCG 480
 QY 481 GGCCATCAACAGGCACTCTGACGACACACGCTGGGACACCGCAGCCGCGCTCGTTCAT 540
 DB 481 GGCCATCAACAGGCACTCTGACGACACACGCTGGGACACCGCAGCCGCGCTCGTTCAT 540
 QY 541 GATCGCGCTGCGCGGCGGTGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGG 600
 DB 541 GATCGCGCTGCGCGGCGGTGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGG 600
 QY 601 GCGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCGGGAACCTCTCTATGC 660
 DB 601 GCGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCGGGAACCTCTCTATGC 660
 QY 661 CGCCTTCTCACCGCGCGCGCTTCCACTGCGCGCTTGGCGGTGGTGGTGTCTACCG 720
 DB 661 CGCCTTCTCACCGCGCGCGCTTCCACTGCGCGCTTGGCGGTGGTGGTGTCTACCG 720
 QY 721 GAAGATCTACGAGCGCGGCAAGTTTCGTTTCGCGCGCGCGCGGAGAGCTGTGTCGCGTT 780
 DB 721 GAAGATCTACGAGCGCGGCAAGTTTCGTTTCGCGCGCGCGCGGAGAGCTGTGTCGCGTT 780
 QY 781 GCGCGCCACCATGCGAGGTGAGGTCCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGT 840
 DB 781 GCGCGCCACCATGCGAGGTGAGGTCCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGT 840
 QY 841 GTTCAGGACATTCGAAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGGA 900
 DB 841 GTTCAGGACATTCGAAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGGA 900
 QY 901 CGAGAAGGAGAGCGGAGCAGCCATGATGTTGGGAATTCGATTGGCGGTGTTTGTCTGTG 960
 DB 901 CGAGAAGGAGAGCGGAGCAGCCATGATGTTGGGAATTCGATTGGCGGTGTTTGTCTGTG 960
 QY 961 CTGGATCCCTTCTTCTGACCGAACTCATCGCCCACTCTGTGCTGAGCGCTGCCCC 1020
 DB 961 CTGGATCCCTTCTTCTGACCGAACTCATCGCCCACTCTGTGCTGAGCGCTGCCCC 1020
 QY 1021 CATCTGGAAGAGCATATTTCTGTGGCTGGCTACTCCAATTTCTTCAACCCCTGAT 1080
 DB 1021 CATCTGGAAGAGCATATTTCTGTGGCTGGCTACTCCAATTTCTTCAACCCCTGAT 1080
 QY 1081 TTACACAGCTTTTAAACAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAG 1140
 DB 1081 TTACACAGCTTTTAAACAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAG 1140
 QY 1141 ATGAACACAGGG 1152
 DB 1141 ATGAACACAGGG 1152

RESULT 2
 ABA01984
 ID ABA01984 standard; cDNA; 1150 BP.
 XX
 AC ABA01984;
 XX
 DT 12-FEB-2002 (first entry)
 XX
 DE Human NOV5a coding sequence.
 XX
 KW Human; NOV5a; metabolic disorder; neurodegenerative disorder;
 KW immune disorder; haematopoietic disorder; developmental disease; cancer;
 KW retinal disease; feeding disorder; vaccine; infection; gene therapy;
 KW neurological disorder; psychotic disorder; G-protein coupled receptor;
 KW cytotatic; antidiabetic; virucide; neuroprotective; nootropic;
 KW analgesic; antidepressant; antimigraine; anticonvulsant; neuroleptic;
 KW antiasthmatic; antiallergic; antiinflammatory; anorectic; antiarthritic;
 KW antipsoriatic; antiatherosclerotic; antibacterial; fungicide;
 KW osteopathic; protozoacide; antiulcer; hypertensive; hypotensive;
 KW antiinfertility; vulnery; nephrotropic; antilipemic;


```

Db 406 CTGTGCTGCCCCGCGCTCGGGAAGTGTCGGGCATCGCCCTGGCGCCGACGGGGCC 465
QY 485 ATCACAGGCACCTGACGACACAGCTGGCAGCCCGGAGCGCGCTCGTTGCTCATGATC 544
Db 466 ATCACAGGCACCTGACGACACAGCTGGCAGCCCGGAGCGCGCTCGTTGCTCATGATC 525
QY 545 GCGCTGCGCCGGGTGCGCTGCGCTCATCGCCCTGCGCCGCTGCTCTTTGGCGGGGC 604
Db 526 GCGCTGCGCCGGGTGCGCTGCGCTCATCGCCCTGCGCCGCTGCTCTTTGGCGGGGC 585
QY 605 GAGGTGTGCGACGCTGGCTCCAGCGCTGCCAGGTGAGCCGGGAACCCCTCTATGCCGCC 664
Db 586 GAGGTGTGCGACGCTGGCTCCAGCGCTGCCAGGTGAGCCGGGAACCCCTCTATGCCGCC 645
QY 665 TTCTCCACCGGGGCGCTTCCACCTGCGCTTGGGTTGGTGGCTTGTCTACCGGAG 724
Db 646 TTCTCCACCGGGGCGCTTCCACCTGCGCTTGGGTTGGTGGCTTGTCTACCGGAG 705
QY 725 ATCTACGAGCGCGCCAAAGTTTCGTTTGGCGCGCGCGGAGAGCTGTGCTGCGCTTGGCG 784
Db 706 ATCTACGAGCGCGCCAAAGTTTCGTTTGGCGCGCGCGGAGAGCTGTGCTGCGCTTGGCG 765
QY 785 GCCACATGACGTGAGTCCAAAGTAAAGGAAGCACCTGTATGAGGCTGAAGTGTGTTC 844
Db 766 GCCACATGACGTG-----AAGGAAGCACCTGTATGAGGCTGAAGTGTGTTC 813
QY 845 AGCGACATGTAAGACAGAGGTGCTTCCAGGTGAGCGGGGACTCCCTGGCGGGAGCAG 904
Db 814 AGCGACATGTAAGACAGAGGTGCTTCCAGGTGAGCGGGGACTCCCTGGCGGGAGCAG 873
QY 905 AAGGAGAGCGAGCGCCATGATGTTGGGAATTCGATTGGCGTGTGTTGCTGTCTGG 964
Db 874 AAGGAGAGCGAGCGCCATGATGTTGGGAATTCGATTGGCGTGTGTTGCTGTCTGG 933
QY 965 ATCCGCTTCTTCTGACGAACTCATPAGCCCACTGTGCTGCGAGCTGCCCCCATC 1024
Db 934 ATCCGCTTCTTCTGACGAACTCATPAGCCCACTGTGCTGCGAGCTGCCCCCATC 993
QY 1025 TGGAAAGCATATTTCTGGCTTGGCTACTCCAAATTCCTTCAACCCCTGATTTAC 1084
Db 994 TGGAAAGCATATTTCTGGCTTGGCTACTCCAAATTCCTTCAACCCCTGATTTAC 1053
QY 1085 ACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1144
Db 1054 ACAGCTTTTAAAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAGATGA 1113
QY 1145 ACACAGGG 1152
Db 1114 ACACAGGG 1121

```

RESULT 5

AAD24963
ID AAD24963 standard; cDNA; 1119 BP.

XX AC AAD24963;

DT 12-MAR-2002 (first entry)

DE Human G-protein coupled receptor-8 (GREC-8) cDNA.

XX Human; G-protein coupled receptor-8; GREC-8; therapy; cancer; stroke;
 KW cell proliferative disorder; neurological; epilepsy; Parkinson's disease;
 KW Alzheimer's disease; inflammation; thyroiditis; haemolytic anaemia; AIDS;
 KW Acquired Immune Deficiency Syndrome; dementia; nontropic; cholelithiasis;
 KW multiple sclerosis; atherosclerosis; angina pectoris; gastroenteritis;
 KW diabetes; ulcer; viral infection; immunosuppressive; ss.

OS Homo sapiens.

XX Key Location/Qualifiers
 FH 1. .1119
 CDS

```

FT /*tag= a
FT /product= "Human GREC-8 protein"
FT 1..195
FT sig_peptide
FT /*tag= b
FT 196..1116
FT mat_peptide
FT /*tag= c
FT /product= "Mature GREC-8 protein"
XX WO200198351-A2.
XX 27-DEC-2001.
XX 15-JUN-2001; 2001WO-US019275.
XX 16-JUN-2000; 2000US-0212483P.
XX 22-JUN-2000; 2000US-0213954P.
XX 29-JUN-2000; 2000US-0215209P.
XX 07-JUL-2000; 2000US-0216595P.
XX 14-JUL-2000; 2000US-0218936P.
XX 19-JUL-2000; 2000US-0219154P.
XX 21-JUL-2000; 2000US-0220141P.
XX (INCY-) INCYTE GENOMICS INC.
XX Lal P, Baughn MR, Hafalia AJA, Nguyen DB, Gandhi AR, Kallick DA,
XX Griffin JA, Yue H, Khan FA, Patterson C, Lu DAM, Tribouley CM;
XX Lu Y, Wallia NK, Graul R, Yao MG, Yang J, Ramkumar J, Au-Young J;
XX Elliott VS, Hernandez R, Walsh RT, Borowsky ML, Thornton M, He A;
XX WPI; 2002-075627/10.
XX P-PSDB; AAE15638.
XX Isolated human G-protein coupled receptor polypeptides and the use of
XX these sequences in the diagnosis, treatment and prevention of diseases
XX and in the assessment of exogenous compounds on the expression of the
XX receptors.
XX Claim 11; Page 136-137; 143pp; English.
XX The invention relates to isolated human G-protein coupled receptor
XX (GREC) polypeptides and their biologically active fragments. GREC and
XX protein is useful in treating a disease or condition associated with an
XX increase or decrease in expression of functional GREC. The GREC's are
XX useful in the diagnosis, treatment and prevention of cell proliferative
XX disorders (cancer, leukaemia, melanoma); neurological disorders (stroke,
XX epilepsy, Parkinson's disease, dementia; Alzheimer's disease); autoimmune
XX inflammatory disorder (thyroiditis, haemolytic anaemia, AIDS, multiple
XX sclerosis); cardiovascular disorder (atherosclerosis, angina pectoris),
XX gastrointestinal disorder (ulcer, cholelithiasis, gastroenteritis),
XX metabolic disorders (diabetes); viral infections (herpes virus) and in
XX the assessment of the effects of exogenous compounds on the expression of
XX the nucleic acid and amino acid sequences. The present sequence is human
XX GREC-8 cDNA
XX
XX Sequence 1119 BP; 171 A; 376 C; 346 G; 226 T; 0 U; 0 Other;

```

Query Match 92.0%; Score 1060; DB 6; Length 1119;

Best Local Similarity 97.7%; Pred. No. 6.1e-220;

Matches 1114; Conservative 0; Mismatches 5; Indels 21; Gaps 3;

QY 5 ATGAGAGCGCTAGCCTTTACGTGGCCACCGCGCGGTTGCCCTTGCCTGGACCCGAG 64

Db 1 ATGAGAGCGCTAGCCTTTACGTGGCCACCGCGCGGTTGCCCTTGCCTGGACCCGAG 60

QY 65 ACACAGCGGACCGCGGACCCCAAGCCGAGAGGATCTCGTTCCACCCGAGCGGC 124

Db 61 ACACAGCG-----CGGACCCCAAGCCGAGAGGATCTCGTTCCACCCGAGCGGC 114

QY 125 GCCGTCTCTGCCGGCGGAGGGCGCCCTTCTCTGTCTTTCACGGTCTCTGTGTGACGCTG 184

Db 115 GCCGTCTCTGCCGGCGGAGGGCGCCCTTCTCTGTCTTTCACGGTCTCTGTGTGACGCTG 174

QY 185 CTAGTGCTGCTGATCGTGCACCTTTCCTGTGTGAACCTGCTGTTCGGTCCACCATCCCG 244

```
Db 175 CTAGTGTCTGATCGCGCACCTTTCTCTGTGGAACCTGCTGTGGTTCACCATCCCG 234
Qy 245 CGGGTCGGTCTTCCACCGGGTCCCGATACCTTGTGGCTCGACGGCGTCTCGGAC 304
Db 235 CGGGTCGGTCTTCCACCGGGTCCCGATACCTTGTGGCTCGACGGCGTCTCGGAC 294
Qy 305 GAACCTAGTGGCAGCGCTTCCACCGGGTCCCGATACCTTGTGGCTCGACGGCGGGA 364
Db 295 GAACCTAGTGGCAGCGCTTCCACCGGGTCCCGATACCTTGTGGCTCGACGGCGGGA 354
Qy 365 CGTGGCTGTGGTGGCGGAGCTTGTGCCACGTGTGGATCTCTTCGACGCCGAGCCGTGT 424
Db 355 CGTGGCTGTGGTGGCGGAGCTTGTGCCACGTGTGGATCTCTTCGACG-----TG 405
Qy 425 CTGTGTCTGGCGCGCGCTCGGAAAGTGGCGGCGCATCGCCCTGGGCGCGGAGCGGGCC 484
Db 406 CTGTGTCTGGCGCGCGCTCGGAAAGTGGCGGCGCATCGCCCTGGGCGCGGAGCGGGCC 465
Qy 485 ATCACAGGCACTTCGACGACACGCTGCGCACCGCGAGCGCGCTCGTTGCTCATGATC 544
Db 466 ATCACAGGCACTTCGACGACACGCTGCGCACCGCGAGCGCGCTCGTTGCTCATGATC 525
Qy 545 GCGCTCGCGCGGGTCCCGTCCAGCGCTCCAGCGCTCCAGCGTCCAGCGTCCAGCGTCCAGCG 604
Db 526 GCGCTCACCGGGTCCCGTCCAGCGCTCCAGCGCTCCAGCGTCCAGCGTCCAGCGTCCAGCG 585
Qy 605 GAGTGTGGCAGCGTCCCGTCCAGCGCTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCG 664
Db 586 GAGTGTGGCAGCGTCCCGTCCAGCGCTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCG 645
Qy 665 TTCTCCACCGCGCGCTTCCAGCGCTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCG 724
Db 646 TTCTCCACCGCGCGCTTCCAGCGCTCCAGCGTCCAGCGTCCAGCGTCCAGCGTCCAGCG 705
Qy 725 ATCTACGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGCGCGAGCTGTGCTCGGTTGCGG 784
Db 706 ATCTACGAGCGCGCAAGTTTCTGTTTGGCGCGCGCGCGCGAGCTGTGCTCGGTTGCGG 765
Qy 785 GCGACCATGAGGTGAGTCCAGGTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 844
Db 766 GCGACCATGAGGTGAGTCCAGGTAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 819
Qy 845 ACGGCACATTCAAGCAAGCAAGTGTCTTCCAGGTGAGCGGGGACTCTCGCGGAGCAG 904
Db 820 ACGGCACATTCAAGCAAGCAAGTGTCTTCCAGGTGAGCGGGGACTCTCGCGGAGCAG 879
Qy 905 AAGGAGAGCGGAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 964
Db 880 AAGGAGAGCGGAGCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 939
Qy 965 ATCCCTCTTCTGAGCGGAACATCATGAGCCCATCTGTGCTCGAGCTCGCCCGCCATC 1024
Db 940 ATCCCTCTTCTGAGCGGAACATCATGAGCCCATCTGTGCTCGAGCTCGCCCGCCATC 999
Qy 1025 TGGAAAGCATATTTCTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1084
Db 1000 TGGAAAGCATATTTCTGTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1059
Qy 1085 ACAGCTTTTAAAGAACTACAAAGTCCCTTAAAGAGCTCTTTTAAAGAGAGATGA 1144
Db 1060 ACAGCTTTTAAAGAACTACAAAGTCCCTTAAAGAGCTCTTTTAAAGAGAGATGA 1119
```

RESULT 6

ABK71923

ID ABK71923 standard; cDNA; 1155 BP.

XX

AC ABK71923;

XX

DT 30-JUL-2002 (first entry)

XX

DE Human cDNA encoding hydroxytryptamine receptor-like protein NOV8.

Human; ss; gene; NOVX; developmental disorder; endocrine disorder; vascular disorder; infectious disease; anorexia; cancer; stroke; neurodegenerative disorder; Alzheimer's disease; acute brain injury; central nervous system disorder; depression; lung disorder; reproductive disorder; tissue disorder; thrombocytopaenia; migraine; angiodysplasia; asthma; X-linked severe combined immunodeficiency; inflammation; autoimmune disorder; immune disorder; blood disorder; haematopoietic disorder; gastrointestinal disease; respiratory disorder; hepatitis; fertility; hypertension; arteriosclerosis; ischaemia; rheumatoid arthritis; Grave's disease; wound healing.

Homo sapiens.

WO200224733-A2.

28-MAR-2002.

17-SEP-2001; 2001WO-US029115.

15-SEP-2000; 2000US-0232675P.

15-SEP-2000; 2000US-0232676P.

15-SEP-2000; 2000US-0232679P.

18-SEP-2000; 2000US-023382P.

18-SEP-2000; 2000US-0233402P.

19-SEP-2000; 2000US-0233521P.

19-SEP-2000; 2000US-0233522P.

19-SEP-2000; 2000US-0233801P.

20-SEP-2000; 2000US-0233960P.

06-OCT-2000; 2000US-0238398P.

13-OCT-2000; 2000US-0240284P.

13-OCT-2000; 2000US-0240498P.

11-JAN-2001; 2001US-0260973P.

26-JAN-2001; 2001US-0264274P.

09-MAR-2001; 2001US-0274862P.

(CURA-) CURAGEN CORP.

Mishra VS, Syptek KA, Taupier RJ, Vernet CAM, Colman SD; Gorman L, Tchernev VT, Malyankar UM, Shenoy S, Tchernev VT; Padigaru M, Patturajan M, Burgess CE, Smithson G, Mallet I; Peyman JA, Stone D, Gunther E, Ellerman K;

WPI; 2002-383182/41.

P-PSDB; ABG60235.

New cytoplasmic, nuclear, membrane bound and secreted NOVX polypeptides, useful for treating cancers and tumors, lung disorders, hematopoietic disorders, autoimmune diseases and immune disorders.

Claim 9; Page 59-60; 210pp; English.

The invention relates to an isolated NOVX polypeptide selected from NOV1a, NOV1b, NOV1c, NOV2a, NOV2b, NOV3a, NOV3b, NOV4a, NOV4b, NOV5a, NOV5b or NOV6-NOV9 polypeptides, their mature form or variant. Also included are a nucleic acid encoding a NOVX protein or variant; a vector comprising the nucleic acid; a cell comprising the vector; an anti-NOVX antibody; and identifying agents that modulate the expression or activity of NOVX. NOVX, the nucleic acid, antibody and modulators are useful in the diagnosis, treatment or prevention of developmental disorders, endocrine disorders, vascular disorders, infectious disease, anorexia, cancer, neurodegenerative disorders (e.g. Alzheimer's disease, Parkinson's disease, Huntington's disease, multiple sclerosis and amyotrophic lateral sclerosis), acute brain injury (e.g. stroke, head injury and cerebral palsy), central nervous system disorders (e.g. depression, epilepsy and schizophrenia), lung disorders, reproductive disorders, disorders affecting carbohydrate metabolism (e.g. galactosaemia and hereditary fructose intolerance), tissue disorders (e.g. Wiskott-aldrich syndrome, thrombocytopaenia, night blindness and sickle cell disease), disorders linked to abnormal angiogenesis, asthma, azoospermia, learning disabilities, facial dysmorphism, autoimmune encephalomyelitis, X-linked severe combined immunodeficiency, seizures, migraines, inflammation, autoimmune disorders, disorders affecting sleep,

CC appetite, thermoregulation, pain, perception, hormone secretion and
CC sexual behaviour, immune disorders, haemopoietic disorders or other
CC disorders related to cell signal processing and metabolic pathway
CC modulation, gastrointestinal diseases, respiratory disorders, blood
CC infections, hepatitis, trauma, regeneration, viral, bacterial or parasitic
CC infections, hyper- or hypo-thyroidism, endometriosis, fertility,
CC hypertension, arteriosclerosis, ischaemia, haemolytic anaemia, Werner
CC syndrome, rheumatoid arthritis, Grave's disease, wound healing, X-linked
CC mental retardation, psychotic and neurological disorders and neuronal
CC degeneration. The present sequence encodes a NOVX protein
XX
SQ Sequence 1155 BP; 175 A; 397 C; 356 G; 227 T; 0 U; 0 Other;

Query Match 88.2%; Score 1016; DB 6; Length 1155;

Best Local Similarity 97.2%; Pred. No. 2.1e-210; Length 1155;

Matches 1087; Conservative 0; Mismatches 0; Indels 31; Gaps 4;

QY 36 CCGGCGTTGCCCTTCCCTGGAGCCCGAGACGAGCAGC-
DB 68 CCGGCGTTGCCCTTCCCTGGAGCCCGAGACGAGCAGCAGGAGCCCGGAGCCCGGCGG 127
QY 95 AGAGGAGTACTCGGTTCGACCCCGAGCGCGCGCTCTCGCGGGCGAGGGCGCCCTTC 154
DB 128 AGAGGAGTACTCGGTTCGACCCCGAGCGCGCGCGCTCTCGCGGGCGAGGGCGCCCTTC 187
QY 155 TCTGTCTTCACGGTCTCGGTGAGCGCTGCTAGTGTCTGATCGCTGCCACTTTCCTG 214
DB 188 TCTGTCTTCACGGTCTCGGTGAGCGCTGCTAGTGTCTGATCGCTGCCACTTTCCTG 247
QY 215 TGGAACTCTGCTGCTCCGATACCATCCCGCGGGTCCGTGCTTCCACCGCGTGCAGAT 274
DB 248 TGGAACTCTGCTGCTCCGATACCATCCCGCGGGTCCGTGCTTCCACCGCGTGCAGAT 307
QY 275 AACTTGGTGGCTCGACGCGCGTCTCGAGCAACTAGTGGAGCGCTGCGATGCCACCG 334
DB 308 AACTTGGTGGCTCGACGCGCGTCTCGAGCAACTAGTGGAGCGCTGCGATGCCACCG 367
QY 335 AGCTTGGGAGTGAAGTCTGACCGCGGCGAGCTGCGTCTGCGCGGAGCGCTGTCAC 394
DB 368 AGCTTGGGAGTGAAGTCTGACCGCGGCGAGCTGCGTCTGCGCGGAGCGCTGTCAC 418
QY 395 GTGTGGATCTCTTCGACCGCGGAGCGCTGCTGTGTGCTGCCCGCGCGCTCGGAGCGTG 454
DB 419 GTGTGGATCTCTTCGAGCGC-----CTGTGCTGCCCGCGCGCTCGGAGCGTG 469
QY 455 GGGGCGATCGCCCTGGCGCGGAGCGGCGCATCACAGGCACTGACAGCAGCGTGGCG 514
DB 470 GCGGCGATCGCCCTGGCGCGGAGCGGCGCATCACAGGCACTGACAGCAGCGTGGCG 529
QY 515 ACCCGGAGCGCGCTCGTGTCTCATGTCGCGCTCGCCCGGCTGCGTGGCGCTCATC 574
DB 530 ACCCGGAGCGCGCTCGTGTCTCATGTCGCGCTCGCCCGGCTGCGTGGCGCTCATC 589
QY 575 GCGCTGGCGCGCTGCTTTGGCGGGCGAGGTGTGCGAGCTCGGCTCGAGCGCTGC 634
DB 590 GCGCTGGCGCGCTGCTTTGGCGGGCGAGGTGTGCGAGCTCGGCTCGAGCGCTGC 649
QY 635 CAGGTGAGCGGGAACCTCTATCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCGG 694
DB 650 CAGGTGAGCGGGAACCTCTATCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCGG 709
QY 695 CTTGCGGTGGTCCGCTTTGTCTACCGGAAGATCTACGAGCGCGCGCAAGTTTCGTTCCGCG 754
DB 710 CTTGCGGTGGTCCGCTTTGTCTACCGGAAGATCTACGAGCGCGCGCAAGTTTCGTTCCGCG 769
QY 755 GCGCGCGGAGAGCTGTGCTGCGCTTTCGCGGCGACCATGACAGGTGCAAGGTAAAG 814
DB 770 GCGCGCGGAGAGCTGTGCTGCGCTTTCGCGGCGACCATGACAGGTG-----AAG 817
QY 815 GAAGACCTGATGAGCGTGAAGTGTGTTACGGGACATTCGAGGACGAGTGTCCCTTC 874
DB 818 GAAGACCTGATGAGCGTGAAGTGTGTTACGGGACATTCGAGGACGAGTGTCCCTTC 877

QY 875 CAGGTGAGCGGAGCTCCTGGCGGAGCAGAGGAGGAGCGAGCAGCCATGATGTGGGA 934
DB 878 CAGGTGAGCGGAGCTCCTGGCGGAGCAGAGGAGGAGGAGCAGCCATGATGTGGGA 937
QY 935 ATTCTGATTGGCGTGTGTTGTGCTGTGATCCCTCTTCTGACGGAACATCATCAGC 994
DB 938 ATTCTGATTGGCGTGTGTTGTGCTGTGATCCCTCTTCTGACGGAACATCATCAGC 997
QY 995 CCACTCTGTGCTGAGCGCTGCCCGCATCTGGAAGCATATTTCTGTGGCTTGGCTAC 1054
DB 998 CCACTCTGTGCTGAGCGCTGCCCGCATCTGGAAGCATATTTCTGTGGCTTGGCTAC 1057
QY 1055 TCCAATTCTTCTTCAACCCCTGATTTACACAGCTTTTAAACAAGAACTACAACATGCC 1114
DB 1058 TCCAATTCTTCTTCAACCCCTGATTTACACAGCTTTTAAACAAGAACTACAACATGCC 1117
QY 1115 TTCAAGAGCGCTCTTTACTAAGCAGAGATGAACACAGGG 1152
DB 1118 TTCAAGAGCGCTCTTTACTAAGCAGAGATGAACACAGGG 1155
RESULT 7
ADC86156
ID ADC86156 standard; DNA; 3086 BP.
XX
AC ADC86156;
XX
DT 01-JAN-2004 (first entry)
XX
DE Human GPCR gene SEQ ID NO:609.
XX
ds; gene; human; GPCR;
KW guanosine triphosphate-binding protein coupled receptor; gene therapy.
XX
OS Homo sapiens.
XX
PN BP1270724-A2.
XX
PD 02-JAN-2003.
XX
PF 18-JUN-2002; 2002EP-00013517.
XX
PR 18-JUN-2001; 2001JP-00246789.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
XX
PI Suwa M, Asai K, Akiyama Y, Aburatani H;
XX
DR WPI; 2003-315783/31.
DR P-PSDB; ADC86157.
XX
PT New polynucleotide, useful for preparing a composition for treating a
PT patient in need of increased or suppressed activity or expression of the
PT guanosine triphosphate-binding protein coupled receptor.
XX
PS Claim 1; SEQ ID NO 609; 28pp; English.
XX
CC The invention relates to a novel polynucleotide encoding a guanosine
CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
CC the invention may have a use in gene therapy. The polynucleotide and
CC polypeptide are useful for preparing a composition for treating a patient
CC in need of increased or suppressed activity or expression of the
CC guanosine triphosphate-binding protein coupled receptor. The
CC polynucleotide sequences shown in ADC85548-ADC87616 encode GPCR's of the
CC invention.
SQ Sequence 3086 BP; 619 A; 885 C; 937 G; 645 T; 0 U; 0 Other;

Query Match

Best Local Similarity 65.6%; Score 756.2; DB 9; Length 3086;

Matches 810; Conservative 0; Mismatches 13; Indels 26; Gaps 2;


```
Db 717 TTGTGTCACCGCCAGCACTGGAACGTGGGGCCATCGCCCTGGATCGTACTGGAAT 776
Qy 485 ATACACGGACACCTGACAGACACGCTGCGACACCGCGCGCTCGTGTCTCATGATC 544
Db 777 ATACCGGGCCACCTGCACTGACGCTGCGACACCGCGAGCGCTGCTTGTGGCTCATGATC 836
Qy 545 GCGCTCGCCGGGTGCGCTGCGGCGCTCATCGCCCTCGCGCGCTGCTTGTGGCGGGC 604
Db 837 GCGATCACCTGGGCACTGTGCGCGCTCATGTCTTCGCGCGCTGCTTGTGGCTGGGC 896
Qy 605 GAGGTGTGCGACGCTCGGCTCCAGCGTGCAGGTGAGCGCGGAAACCTCTCATGCGGCC 664
Db 897 GAACCTATGATGCTCGGCTGCAGCGTTCAGAGTGAGCCAGAGCCCTCTCATGCTGC 956
Qy 665 TTCTCCACCGCGCGGCTTCACACTGCGCGTTCGCGGTGGCGGTGCTGCTTGTACCGAAG 724
Db 957 TTCTCCACCTGCGGAGCGCTTACCTGCTCTAGCGGTGGTGTCTTGTCTACTGGAAA 1016
Qy 725 ATCTACGAGCGCGCAAGTTTCTGTTTCGCGCGCGCGCGAGAGCTGTGCTGCGGTGCG 784
Db 1017 ATATACAAAGCCGCCAAGTTTCGATTTCGTCGACAGCGCGCGGTGGTACCGCTTCT 1076
Qy 785 GCCACCATGCAAGTGAGTCCAAAGTAAAGGAAGCACCTGATGAGGCTGAAGTGTGTC 844
Db 1077 GCCACCAAGC-----AGCAAGGAAGCACCTCCGGAGTCTGAGATGTTGTC 1124
Qy 845 ACGGCAATGCAAGCAACCGTGTCTTCCAGGTGAGCGGGGACTCTGCGGGGAGAG 904
Db 1125 ACAGCCGCTGCGCGAGCAACAGTGACCTTCCAGACAAGCGGAGACTCTGCGGGGAGCAG 1184
Qy 905 AAGGAGGCGGAGCAGCATGATGTTGGGAACTCTGATTGGCGGTGTTGCTGCTGCTGG 964
Db 1185 AAGGAGAAGCGGCGAGCAGATGTTGGGATCTTGAATGTCGCGGTGTTGCTTGTGG 1244
Qy 965 ATCCCTCTTCTTCGAGGAACTCATCAGCCCACTCTGTGCTTCGAGCTGCCCCCAATC 1024
Db 1245 ATCCCTCTTCTTCGAGGAGCTCATCAGCCCGCTCTGTGCTGCAGGCTGCCACCAATC 1304
Qy 1025 TGGAAAAGCATATTTCTGTGCTGGTACTCCAAATTTCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAAGCATATTTCTGTGCTGGATATTCCAATTCGTTCTTCAACCCCTGATTTAC 1364
Qy 1085 ACAGCTTTTACAGAACTACAAGTCCCTTCAGAGCTCTTACTAAGCAGAGATGA 1144
Db 1365 ACTGCTTTAAGATTAACAAGTCCCTTCAGAGCTCTTACTAAGCAGAGATGA 1424

RESULT 10
RAD48747
XX AAD48747 standard; DNA; 2061 BP.
XX AC
XX AAD48747;
DT 07-MAR-2003 (first entry)
XX Mouse 5-HT5B receptor gene.
XX
KW Mouse; 5-hydroxytryptophan; 5-HT5B receptor; depression; transgenic;
KW transgenic animal; phenotype; pain sensitivity; neurological disease;
KW gene therapy; pharmacological; neuropsychological disease; analgesic;
KW psychotic illness; receptor; gene; ds.
XX
OS Mus sp.
XX
XX Key Location/Qualifiers
XX misc_feature 72..391
XX /tag= a
XX /note= "Sequence flanking Neo insert in targeting
XX construct"
XX 312..1424
XX /tag= b
XX /product= "Mouse 5-HT5B receptor protein"
```

```
FT misc_feature 392..464
FT /tag= C
FT /note= "Sequence deleted in targeting construct"
FT misc_feature 465..1008
FT /tag= d
FT /note= "Sequence flanking Neo insert in targeting
FT construct"
XX
XX WO200279443-A2.
XX
XX 10-OCT-2002.
XX
XX 29-MAR-2002; 2002WO-US009853.
XX
XX 29-MAR-2001; 2001US-0280553P.
XX
XX 21-DEC-2001; 2001US-0342472P.
XX
XX 28-MAR-2002; 2002US-00109532.
XX
XX (DELT-) DELTAGEN INC.
XX
XX Allen KD;
XX
XX WPI; 2003-067450/06.
XX
XX P-PSDB; AAE48747.
XX
XX Novel transgenic mouse useful for identifying a potential therapeutic
XX agent for the treatment of depression or pain, comprises disruption in a
XX 5-hydroxytryptophan 5B receptor gene.
XX
XX Example 1; Fig 1; 60pp; English.
XX
XX The present invention relates to transgenic animals, compositions and
XX methods relating to the characterisation of gene function. The invention
XX also relates to transgenic mice comprising mutations in 5-hydroxy-
XX tryptophan (5-HT)5B receptor gene. The transgenic mice are useful for
XX identifying an agent that modulates the phenotype such as increased
XX depression or increased pain sensitivity. They are also useful for
XX identifying potential therapeutic agents for the treatment of pain or
XX depression and for evaluating a potential therapeutic agent capable of
XX affecting a condition associated with a mutation in a 5-HT5B receptor
XX gene. Transgenic animals of the invention are also useful for testing the
XX efficacy of proposed genetic and pharmacological therapies for human
XX diseases such as neurological, neuropsychological or psychotic illnesses.
XX The present sequence is mouse 5-HT5B receptor gene
XX
XX Sequence 2061 BP; 446 A; 582 C; 545 G; 488 T; 0 U; 0 Other;
```

```
Query Match 63.6%; Score 732.8; DB 7; Length 2061;
Best Local Similarity 79.9%; Pred. No. 4.9e-149;
Matches 911; Conservative 0; Mismatches 202; Indels 27; Gaps 3;

Qy 5 ATGGAGCGCGTAGCCTTTCAGTGGCCACCGCGCGTTCCTTGCCTCGGCCCGGAG 64
Db 312 ATGGAAGTTTCTAACTCTCAGCGCGCACTCCCGCCCTTGCCTTCTCCCGGAGCTGAG 371
Qy 65 ACCAGCAGCGAGCCCGGAGCCCAAGCCCGAGAGGATCTCGGTTTCACCCCGGAGCGGC 124
Db 372 AGCTGCAGTGAC-----AGCCCAAGTTCGCGCAGGAGCATGGGATCCACCCAGTGGG 425
Qy 125 GCCGTCTTCGCGGCGGAGGCGCCCTTCTCTGTCTTACGGTCTCTGTGTGTGAGCTG 184
Db 426 CTCATCTTTCGCGCGCGGAGCGCCCTTCTCTGTCTTTCACGCTGCTTGTGTGACTCTA 485
Qy 185 CTAGTGTGCTGATGCTGCGCACTTCTCTGTGGAACTCTGCTGTGCTGCTGCTGCTGCT 244
Db 486 CTGCTGTGCTGATGCTGCGCACTTCTTATGGAATCTGCTAGTCTGCTGCTGCTGCTGCT 545
Qy 245 CGGCTGCTGCTTCCACCGCGTCCGCAATAACTTGTGTGGCTCGACCGCGCTCTCGGAC 304
Db 546 CGGTCGCGCGCTTCCACCGGTCGCCATTAACCTTGTGTGGCTCGACACCGCTCTCGAT 605
Qy 305 GAACCTAGTGGAGCGCTGCGGATGCCACCGAGCGCTGGCGAGTGTGAGCTGTGAGCCGCGA 364
```



```
Db 606 GTCTGCTGGCGGCTTCTGTGTGATGCTCTAGAGCTGTGTGAGCGAGTGTTCGCTGGCGGA 665
QY 365 CBTGCGCTCTGTGGCGGAGCCCTGTGCCACGTGTGATCTCCTTCGACGCGGAGCCTGT 424
Db 666 CGTTGGCAGTAGCAGAGAGTCTGTGCCACGTGTGATCTCCTTCGACG-----TG 716
QY 425 CTGTGCTGCCCGCGGCTCTGGGAACGTGTGGGGCCATCGCCTGTGGCGCGGAGCGGCC 484
Db 717 TTGTGCTGCACGCCAGCATCTGGAACGTGGGGCCATCGCCTGTGATCGTACTGGACT 776
QY 485 ATCACAGGCACCTGCAGCACAGCTGCGCACCGGACCGGAGCGGCGCTGTGTCTATGATC 544
Db 777 ATACAGCGGCACCTGCAGTACAGCTGCGCACCGGAGCGGCTGTCTCGCTCATGATC 836
QY 545 GCGCTCGCCCGGCTGCGGCGCTCATGCGCCTCGCGCGCTGTCTTTGGCGGGGC 604
Db 837 GCGATCACCTGGGCACTGTCCGGGCTCATGTCTCGCGCGCTGTCTTTGGCTGGGGC 896
QY 605 GAGGTGTGCGAGCTCGGCTCGAGCGCTGCCAGGTGAGCGCGGAACCTCTCTATGCGCC 664
Db 897 GAAGCCTATGATGCTCGGCTGAGCGGTGCCAGGTGAGCGAGCGCCTCCTATGCTGTC 956
QY 665 TTCTCCACCGCGGCGCTTCCACCTGCGCTTGGCGTGGTGCCTTTCTTACCGGAG 724
Db 957 TTCTCCACCTGCGGAGCCTTCTACCTGCTTAGCGGTGGTGTCTCTGCTACTGGAAA 1016
QY 725 ATCTAGAGGCGGCCAAAGTTTCGTTTCGGCGCGCGGAGAGCTGTGCTGCGTTCGCG 784
Db 1017 ATATACAAAGCGCCAAAGTTTCGATTCGTCGACAGCGCGGCGGTGTACGCTTCT 1076
QY 785 GCACATGCGAGGTGAGTCCAAAGTAAAGGAGCACTGTATGAGGCTGAAGTGGTTC 844
Db 1077 GCCACGACG-----AGGCAAGGAGGACCACTCCGAGTCTGAGATGGTGTTC 1124
QY 845 AGCGCACATTCGAAGCAAGCTGTCTTCAGAGTGCAGCGGAGCTCTGCGGGAGCAG 904
Db 1125 ACAGCCGTCGCGGAGCAAGTACAGTACCTTCACAGAGCGGAGCTCTGCGGGAGCAG 1184
QY 905 AAGGAGGCGGAGCAGCCATGATGTGGAAATCTGATTTGGCTGTGTGTGCTGCTGG 964
Db 1185 AAGGAGAGCGGCGCAGCATGATGTCGGATCTTGATTTGGCTGTGTGTGTTGG 1244
QY 965 ATCCCTCTTCTGAGGAACTCATGAGCCCACTGTGCTGCTGAGCCTGCCCCCATC 1024
Db 1245 ATCCCTCTTCTGAGGAGCTCATGAGCCGCTCTGTGCTGAGCCTGCGCCCATC 1304
QY 1025 TGGAAAGCATATTTCTGTGGCTTGTGCTACTCCAAATCTTCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAGCATATTTCTGTGGCTTGTGATATTCATATTCGCTTCAACCCCTGATTTAC 1364
QY 1085 ACAGCTTTTAAAGAACTTAAACAATGCTTTCAAGAGCCTCTTTACTAAGCAGAGATGA 1144
Db 1365 ACTGCTTTTAAAGAACTTAAACAATGCTTTCAAGAGCCTCTTTACTAAGCAGAGATGA 1424
```

RESULT 11

AAQ72270
ID AAQ72270 standard; cDNA; 2226 BP.

XX AAQ72270;

XX 25-MAR-2003 (revised)

DT 24-MAY-1995 (first entry)

XX

DE Rat MR22 serotonin receptor gene.

XX Serotonin; receptor; transmembrane; domain; kinase; phosphorylation;
KW sensory; motor; behaviour; central nervous system; CNS; superfamily;
KW G-protein; ligand-gated; ion channel; subfamily; human; rat; amplify;
KW primer; PCR; amplification; brain; hypothalamus; indolamine; drug;
KW hypothalamus; therapeutic; neurological; pathology; dementia; insomnia;
KW Parkinson's disease; eating disorder; anxiety; migraine; headache; ss.

OS Rattus rattus.

XX Key Location/Qualifiers
CDS 303..1415
FT /tag= a
FT /product= "rat MR22 serotonin receptor"

XX WO9421670-A1.

XX 29-SEP-1994.

XX 15-MAR-1994; 94WO-US002839.

XX 15-MAR-1993; 93US-00031538.

XX (SCRI) SCRIPPS RES INST.

XX Sutcliffe JG, Erlander MG, Lovenberg TW;

XX WPI; 1994-316932/39.

XX P-PSDB; AAR58686.

XX New serotonin receptors and corresp. DNA and antibodies - useful in
diagnosis and treatment of neurological processes and pathologies.

XX Claim 13; Page 137-138; 198pp; English.

XX The nucleotide sequence of the novel rat serotonin receptor MR22. The
gene encodes a protein of 370 amino acids. The protein contains 7
transmembrane domains (TMDs), a putative N-linked glycosylation site and
4 putative sites for kinase C phosphorylation. Serotonin regulates a wide
range of sensory, motor and behavioural functions in the central nervous
system (CNS). Serotonin receptors belong to at least two protein
superfamilies: G-protein-associated receptors containing 7 TMDs
(including 5-HT1a/b/c/d/e, 5-HT2) and ligand-gated ion channel receptors
with 4 TMDs (5-HT3). The serotonin receptors presented in the patent
represent prototypes that fall into three new serotonin subfamily
classifications: 5-HT1e-like (rat MR77 - AAQ72271 and human MR77 -
AAQ72272), 5-HT5 (subdivided into 5-HT5alpha, rat REC17 - AAQ72269 and 5-
HT5beta, rat MR22 - AAQ72270) and 5-HT6 (rat RBC20 - AAQ72273). The genes
for the receptors were cloned by amplification based on conserved amino
acid sequences found in the TMDs. Degenerate primers were constructed to
these sequences and used in two rounds of nested PCR amplification on a
rat brain hypothalamic cDNA template. In the second round of
amplification, degenerate primers corresponding to conserved residues
found only in indolamine-binding receptors were used to amplify only
those receptors. The resultant fragments were cloned into pBluescript
vectors. The plasmid fragments were used to screen a rat hypothalamus
cDNA library to obtain full length clones. The nucleic acids molecules
and their corresponding proteins may be used in methods for determining
ligand binding activity, detecting and altering expression of serotonin
receptors, drug screening and therapeutic treatments involving human
serotonin receptors. Serotonin receptors are also known to participate in
neurological processes. Pathologies including dementia, Parkinson's
disease, eating disorders, pathological anxiety, migraine, headaches,
insomnia and other conditions. (Updated on 25-MAR-2003 to correct PN
field.)

XX Sequence 2226 BP; 488 A; 641 C; 592 G; 505 T; 0 U; 0 Other;

XX Query Match 63.4%; Score 729.8; DB 2; Length 2226;

XX Best Local Similarity 79.6%; Pred. No. 2.2e-148;

XX Matches 911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;

QY 3 CCATGGAGCGCTAGCCTTTTCAGTGGCGGCGGCTTCCCTTCCCTCGGACCGG 62

Db 301 CAATGGAGTCTCTAACTCTCAGCGCCACCTTGGCATTCCTTCTCGGACCGG 360

QY 63 AGACCAGCAGCGGACCCGGGACCCCAAGCCGAGAGGATATCTCGTTTCGACCCCGAGCG 122

Db 361 AGAGCTGCACTGAC-----AGCCCAAGTTCGGCAGAGCATCCACCCAGGGG 414

QY 123 GCGCGTCTCTGCGGCGGCGGCGGCGCTTCTTGTCTTCTACGCTCTGTGGTGAACG 182

Db 415 GGCATCATCTGTTCGGCGCGAGCGCCCTCTCTGCTTACCGTACTCGTGTACTC 474
Qy 183 TGCTAGTGTGCTGATCGCTGCACCTTCTGTTGGAACCTGCTGTTCCGGTACCAATCC 242
Db 475 TACTGGTGTGCTGATCGCTGCCACTTCTTATGGAATCTGCTAGTCTGCTGACTATCC 534
Qy 243 CGGGGTCCGTGCTTCCACGGGTGCGCATACCTTGTGGCTCGACGGCGTCTCGG 302
Db 535 TGGCGGTCCCGGCTTCCACCGGTGCGCATACCTTGTGGCTCGACGGCGTCTCGG 594
Qy 303 ACCAACTAGTGGAGCGCTGCGATGCCAGCGCTGCGGAGTGAAGCTGTGCAACCGGGC 362
Db 595 ACGTCTGTGTGGCGCTCTGGTATGCCACTGAGCTGTGAGGAGTGTGCGCTGGC 654
Qy 363 GAGCTGGCTGTGGCGCGAGCTGTGCGACGTTGTAATCTCTTCGACCGCGAGCCT 422
Db 655 GAGCTGGCAGCTGGCGAGGAGTCTGTGCCACGTGTGATCTCTCTCGAGC----- 705
Qy 423 GTCTGTGCTGCCCGCGGCTCGGAACGTGGCGGCATCGCCCTCGCGCGAGCGGG 482
Db 706 TGTGTGCTGACAGCCAGCATCTGGAACGTGGCGGCATCGCCCTCGATCGCTACTGGA 765
Qy 483 CCATCACAGCGCACTGACGACACAGCTGTGGCACCCCGAGCGCGCTCTGTTGCTCATGA 542
Db 766 CTATCACGCGCCACTGACGTACACGTGCGCACCCCGCGCGCTTACGACTTATGA 825
Qy 543 TCGCGCTGGCGCGGTGCGGTGCGGTCTATGCGCCCTCGCGCGCTGCTTCTTGGCGGG 602
Db 826 TCGCGATCATCTGGGCATCTGTCGCCCTCTATCGCGCTCGCCCGCTGCTTCTTGGCTGG 885
Qy 603 GCGAGTGTGCGAGCTGCGCTCGAGCGCTGCGAGGTGAGCGGGAACCTCTATCGCG 662
Db 886 GCGAAGCCTACGATGCTGCGGTGCGAGCTTCCAGAGTGAGCCAGAGCCTTCGTACGCG 945
Qy 663 CTTCTCTCACCGCGCGGCTTCCACCTGCGCGCTTGGCGTGTGCGCGCTTCTTACCGGA 722
Db 946 TCTTCTCACCTGCGAGCTTCTACGTGCTCTGCGCGTGTGCTCTTCTGCTACTGGA 1005
Qy 723 AGATCTACGAGCGCGCAAGTTCTGTTGCGCGCGCGCGAGAGCTGTGCTGCGGTTC 782
Db 1006 AGATATACAAAGCGCAAGTTTCGATTCGCGCGCGAGAGCGCGGTAGTGCCTCTC 1065
Qy 783 CGGCCACATCGATGAGTCCAGGTAAAGGAGGACCTGATGAGGCTGAAGTGTGT 842
Db 1066 CCGCACACACGC-----AGCAAGGAGGACCTTCAGAGTCTGAGACGGTAT 1113
Qy 843 TCAGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCTGGCGGAGC 902
Db 1114 TCAGCGCGGTTCAGAGCGACAGTGGCTTCCAGACAAGTGGAGACTCTCTGGCGGAGC 1173
Qy 903 AGAAGGAGCGGAGCGACGCCATGATGTTGGGAAATCTGATGGCGTGTGTTGCTGTGCT 962
Db 1174 AGAAGGAGGAGCGCGCCATGATGTTGGGATCTTGTATCGGTGTGTTGCTGTGCT 1233
Qy 963 GGATCCCTCTCTCTGAGGAACTCATCAGCCACTCTGCTGCGAGCTGCGCCCA 1022
Db 1234 GGATCCCTCTCTCTGAGGAGCTCTGCGCCGCTCTGCGCCGAGCTGCGCCCA 1293
Qy 1023 TCTGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTCTTCAACCCCTGATTT 1082
Db 1294 TCTGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTCTTCAATCCCTTAATCT 1353
Qy 1083 ACACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCTCTTTTAAAGCAGAT 1142
Db 1354 ACAGGCTTTTAAAGAACTACAAATGCTTCAAGAGCTCTTTTAAAGCAGAT 1413
Qy 1143 GAACA 1147
Db 1414 AAGCA 1418

ADBS2419
ID ADB52419 standard; DNA; 2240 BP.
XX
AC ADB52419;
XX
DT 04-DEC-2003 (first entry)
XX
DE Primary rat hepatocyte toxicity modelling related gene SEQ ID NO:2961.
XX
KW toxic effect; gene expression profile; hepatotoxicity; diagnostic marker;
KW toxicity marker; toxicity progression; drug screening;
KW primary rat hepatocyte toxicity modelling; gene; ds.
XX
OS Rattus norvegicus.
XX
EN WO2003065993-A2.
XX
PD 14-AUG-2003.
XX
PF 04-FEB-2003; 2003WO-US003482.
XX
PR 04-FEB-2002; 2002US-0353171P.
PR 13-MAR-2002; 2002US-0363534P.
PR 08-APR-2002; 2002US-0370248P.
PR 10-APR-2002; 2002US-0371134P.
PR 10-APR-2002; 2002US-0371135P.
PR 10-APR-2002; 2002US-0371150P.
PR 11-APR-2002; 2002US-0371413P.
PR 19-APR-2002; 2002US-0373601P.
PR 19-APR-2002; 2002US-0373602P.
PR 22-APR-2002; 2002US-0374139P.
PR 08-MAY-2002; 2002US-0378370P.
PR 09-MAY-2002; 2002US-0378652P.
PR 09-MAY-2002; 2002US-0378653P.
PR 09-MAY-2002; 2002US-0378665P.
PR 09-JUL-2002; 2002US-0394230P.
PR 09-JUL-2002; 2002US-0394253P.
PR 04-SEP-2002; 2002US-0407688P.
PR 28-JAN-2003; 2003US-0442900P.
XX
(GENE-) GENE LOGIC INC.
XX
PI Mendrick D, Porter M, Johnson K, Higgs B, Castle A, Orr M;
PI Elashoff M;
XX
XX WPI; 2003-731472/69.
XX
Determining if a compound induces a toxic effect on a tissue or cell, for
identifying hepatotoxic compounds, comprises comparing a gene expression
profile of a tissue or cell sample to a database of Tox mean and non-Tox
mean values.
Claim 44; SEQ ID NO 2961; 874pp; English.
The present invention describes a method for determining whether a
compound induces a toxic effect on a tissue or cell. The method comprises
preparing a gene expression profile of a tissue or cell sample exposed to
the compound, and comparing the gene expression profile to a database
comprising data or information on the Tox mean and non-Tox mean value.
The method is useful for predicting or identifying at least one toxic
effect, particularly hepatotoxicity, of a test or unknown compound. The
genes listed in the specification are useful as diagnostic or toxicity
markers for the prediction or identification of the physiological state
of tissue or cell sample that has been exposed to a compound, or to
identify or predict the toxic effects of a compound or an agent. These
genes may also be used as markers for monitoring toxicity progression or for
drug screening. The present sequence represents a primary rat hepatocyte
toxicity modelling related gene sequence from the present invention.
XX
SQ Sequence 2240 BP; 490 A; 646 C; 597 G; 507 T; 0 U; 0 Other;
Query Match 63.4%; Score 729.8; DB 9; Length 2240;
Best Local Similarity 79.6%; Pred. No. 2.2e-148;

Matches	911;	Conservative	0;	Mismatches	207;	Indels	27;	Gaps	37;
Qy	3	CCATGAGCGCGCTAGCCTTT	CAGTGGCCACCGCGCGTTGCCCTTGCCCTCGGACCCG	62					
Db	301	CAATGGAAGTCTCTAACTCT	CACGGCGCACCCCTGGCATTTGCCCTTCGGGACCCG	360					
Qy	63	AGACGAGCGGACCCGCGGAC	CCCAAGCCGAGGAGATACTCGGTT	TCGACCCCGAGCG	122				
Db	361	AGAGCTGCAGTGAC-----	AGGCCAAAGTTCGGCGAGAAGCATGGAT	CCACCCACAGGG	414				
Qy	123	CGCGCTCCTTGC	CGGGCGGCGGCCCTTCTCTGTCTT	CACGGTCTCGTGTGACGC	182				
Db	415	GGCTCATCTTGT	TCGGCGCGGAGCGGCCCTTCTCTGGCTT	CACCGTACTCGTGTAACTC	474				
Qy	183	TGCTAGTGTGCTGATCGCT	GGCCACTTTCCTGTGGAACCTGCTGGT	TCGGGTCAACATCC	242				
Db	475	TACTGGTGTTCGTGATCGCT	GGCACTTTCTTATGGAATCTGCTAGT	TTCTGTGTACTATCC	534				
Qy	243	CGGGTGTGTCCTTCCACCG	GTGGCGGATAACTTGTGTGGCTTCGA	GGCGCGTCTGG	302				
Db	535	TGGCGTCCGCGCTTCCACCG	TGTGCACATAACTTGGTAGCTTCCACCGCGCTCTCG	594					
Qy	303	ACGAACTAGTGGCAGCGCTG	CGATGCCACCGAGCCTGGCGAGTGAGCT	TCGACCGGGC	362				
Db	595	AGTCTTGTGTGGCGCTCTG	TGTGATGGCACTGAGCCTGGTGGCGAGTGTGCGCTCGGC	654					
Qy	363	GAGCTCGGCTGTGGCGGAG	CGCTGTGCCACGTGTGGATCTCTCT	TCGACCGGAGCCT	422				
Db	655	GACGTTGGCAGCTGGG	CAGAGTCTGTGCCACGTGTGGATCTCTT	CGACG-----	705				
Qy	423	GTCTGTGCTGCTCGCGCG	CGCTCGGAAACGTGGCGGCATCGCCCT	CGGCGCGACGGG	482				
Db	706	TGTTGTCTGCA	CAGCCAGCATCTGGAACGTGGCGGCATCGCCCT	TGGATCGCTACTGGA	765				
Qy	483	CCATCACCGCACCTTCGAC	CACGCTGGCACCCGACGGCGCCTCGT	TGCTATGA	542				
Db	766	CTATCAGCGCCCACTTCAG	TACGCTGGGCACCGCGCGCGCTTCAGCACTTATGA	825					
Qy	543	TCGCGCTCGCGCGGTCCG	CGCTCATCGCCCTCGCGCCTCTCTTTGGCGCGG	602					
Db	826	TCGGGATCACTGGG	CATGTGCGCCCTCATCGGCTTCGCCCGCTCTTTTGGCTGGG	885					
Qy	603	GCAGGTGTGCGACGTTCG	CTCCAGCGCTGCCAGGTGAGCCGGAAACCCCT	CTCATGCCG	662				
Db	886	GCGAAGCCTACGATGCTCG	CTCGCTGCGAGTTCACAGGTGAGCCAGGACCTTCGTACGCCG	945					
Qy	663	CCCTTCTCCACCGCGGCG	CTTCGACCTGCGCTTGGCGTGTGGCGTGTGCTTACCGGA	722					
Db	946	TCCTTCTCCACCTCGG	AGCTTTCTACGTGCTCTGGCGCTGTGCTCTACTGGA	1005					
Qy	723	AGATCTACGAGGCGGCCA	AGTTTCGTTTCGGCGCGCGGAGAGCTGTGCTGCCGTTCG	782					
Db	1006	AGATATACAAGCGCGCA	AGTTTCGATTTCGGCTGACAGCGGGCGGTAGTGCCCTTCG	1065					
Qy	783	CGGCCACCATGCAAGGTG	AGGTCCAAAGTAAGGAAGCACTGATGAGGCTGAAAGTGCTGT	842					
Db	1066	CCGCCACCAAGC-----	AGGCCAAAGAACCACTTCAGGAGTCTGAGACGAT	1113					
Qy	843	TCACGGCATATTCGAAG	CAACGGTGTCTTCCAGGTGACGGGGACTCTCGCGGGAGC	902					
Db	1114	TCACCGCGGTTCG	CAGCGACAGTGCCCTTCACAGCAAGTGGAGACTCTCGCGGGAGC	1173					
Qy	903	AGAAGGAGAGGCGAG	CAGCATCATGTGGGAATCTGATTGGCGTGTGTGCTGTGCT	962					
Db	1174	AGNAGGAGAGCGCG	CACTGATGGTGGGATCTTGATCGGTGTGTGTGTGCTGTGCT	1233					
Qy	963	GGATCCCTCTTCTCTGA	CGGAACCTCATCAGGCCACTCTGTGCTCGCAGCTTCGCCCTCC	1022					
Db	1234	GGATCCCTCTTCTCTGA	CGGAGCTCGTCAGCCGCTCTGCGCTTCGAGCTTCGCCCA	1293					
Qy	1023	TCGTGGAAGACATATTT	CTGTGGCTTGGCTACTCCAAATCTTCTTCTTCAACCCCTCATTT	1082					
Db	1294	TCGTGGAAGACATATTT	CTGTGGCTTGGCTATTCAAATCTGTTCTTCAATCCCTTAACTCT	1353					

QY	1083	ACACAGCTTTTAAACAAGACTACAAACAATGCTTCAAGAGCGCTCTTTACTTAAGCAGAGAT	1144
Db	1354	ACAGCGCTTTTAAAGAAGCTTACAAACAATGCTTCAAGAGCGCTCTTTACTTAAGCAGAGAT	1413
QY	1143	GAACA 1147	
Db	1414	AAGCA 1418	
RESULT 13			
QY	145	AA98145	
ID	AA98145	standard; DNA; 1090 BP.	
XX	AA98145		
AC	AA98145		
XX	AA98145		
DT	12-MAR-2002	(first entry)	
XX	Human DNA	for potential G protein-coupled receptor #102.	
DE	Human; G protein-coupled receptor; GPCR; ds; GAL4; galanin receptor;		
XX	Alzheimer's disease; amyotrophic lateral sclerosis; asthma;		
KW	atherosclerosis; basal cell carcinoma; breast carcinoma; cardiomyopathy;		
KW	chondrosarcoma; chronic obstructive pulmonary disease; Crohn's disease;		
KW	depression; epilepsy; macular degeneration; lymphoma; melanoma;		
KW	multiple sclerosis; osteoarthritis; osteoporosis; Parkinson's disease;		
KW	psoriasis; rheumatoid arthritis; schizophrenia; ulcerative colitis;		
KW	tuberculosis; cognition disorder; memory disorder; anorexia;		
KW	hormonal release disorder; cardiovascular activity disorder;		
KW	pain perception disorder; obesity; diabetes; obesity; diabetes;		
KW	hyperlipidaemia; stroke; gene therapy.		
XX	Homo sapiens.		
OS	WO200185791-A1.		
XX	15-NOV-2001.		
PN	11-MAY-2001; 2001WO-US015332.		
XX	11-MAY-2000; 2000US-0203217P.		
XX	18-MAY-2000; 2000US-0205945P.		
PR	(LIFE-) LIFESPAN BIOSCIENCES INC.		
PR	Brown JP, Miller M, Burmer G, Fabre-Suver C, Pritchard D;		
XX	WPI; 2002-066595/09.		
PI	Novel G protein-coupled receptor polypeptides including galanin receptor		
XX	polypeptides useful for identifying modulators that are useful for		
XX	treating Alzheimer's disease, psoriasis, melanoma, multiple sclerosis,		
XX	stroke.		
PS	Claim 2; Page 103; 144pp; English.		
CC	The invention relates to an isolated polypeptide encoded by a nucleic		
CC	acid molecule that is at least 80% identical to the G protein-coupled		
CC	(GPCR) polynucleotides included in the specification. Also included are		
CC	probes based on the GPCR sequences (including antisense probes), a host		
CC	cell comprising an expression vector comprising the GPCR sequence,		
CC	antibodies raised against the polypeptides, and methods of identifying		
CC	modulators of the polypeptides. The polypeptides are useful for		
CC	identifying modulator compounds which function as modulators, activators,		
CC	repressors, agonists or antagonists of the novel GPCR polypeptides		
CC	including the GAL4 polypeptide. The antibodies and nucleic acid probes as		
CC	described above can be used to detect the presence of the polypeptides		
CC	and nucleic acids and are used to diagnose a variety of diseases or		
CC	disorders in which GPCRs are involved e.g., Alzheimer's disease,		
CC	amyotrophic lateral sclerosis, asthma, atherosclerosis, basal cell		
CC	carcinoma, breast carcinoma, cardiomyopathy, chondrosarcoma, chronic		
CC	obstructive pulmonary disease, Crohn's disease, depression, epilepsy,		
CC	macular degeneration, lymphoma, melanoma, multiple sclerosis,		

CC osteoarthritis, osteoporosis, Parkinson's disease, psoriasis, rheumatoid
CC arthritis, schizophrenia, ulcerative colitis, tuberculosis and many other
CC diseases listed in the specification. The probes and antibodies are also
CC useful for diagnosing cognition and memory disorders, anorexia, hormonal
CC release disorders, cardiovascular activity disorders, pain perception
CC disorders, obesity, diabetes, Alzheimer's disease. Preferably, compounds
CC that decrease or increase the expression of galanin receptor (GAL4) can
CC be used to treat obesity, diabetes, hyperlipidaemia and stroke. The GPCR
CC nucleic acid is useful for treating the above mentioned disorders by gene
CC therapy techniques. The present sequence is a novel GPCR polynucleotide
CC of the invention
XX

QY Sequence 1090 BP; 157 A; 364 C; 372 G; 197 T; 0 U; 0 Other;

Query Match 62.6%; Score 721.2; DB 6; Length 1090;

Best Local Similarity 95.2%; Pred. No. 1.4e-146;

Matches 775; Conservative 0; Mismatches 13; Indels 26; Gaps 2;

QY 36 CCGCGCTTGCCTTGCCTGGACCCGAGACCCGAGACGAGC-GGACCCGGGACCCCAAGCCCG 94
Db 1 CCGCGCTTGCCTTGCCTGGACCCGAGACCCGAGACGAGC-GGACCCGGGACCCCAAGCCCG 60
QY 95 AGAGGATACCTCGTTGACACCCGAGCGCCGCTTCCTGCGGCGGAGGCGCCCTTC 154
Db 61 AGAGGATACCTCGTTGACACCCGAGCGCCGCTTCCTGCGGCGGAGGCGCCCTTC 120
QY 155 TCTGTCTTACGCTTGGTGGTGAAGCTGTAGTGTCTGTGCTGTGCTGCTGCTTCTCG 214
Db 121 TCTGTCTTACGCTTGGTGGTGAAGCTGTAGTGTCTGTGCTGTGCTGCTGCTTCTCG 180
QY 215 TGGAACTGTGCTTCCGCTCACCATCCCGGGGTCCGCTCTCCACCGGTGCGCGAT 274
Db 181 TGGAACTGTGCTTCCGCTCACCATCCCGGGGTCCGCTCTCCACCGGTGCGCGAT 240
QY 275 AACTTGTGGCTCGAGCGCCGCTTCGGAACAACTAGTGGCAGCGCTGGCGATGCCACG 334
Db 241 AACTTGTGGCTCGAGCGCCGCTTCGGAACAACTAGTGGCAGCGCTGGCGATGCCACG 300
QY 335 AGCTGGGAGTGAAGTGTCCACCGGGGAGTGGCTGCTGGCGGAGCGCTGTGCCAC 394
Db 301 AGCTGGGAGTGAAGTGTCCACCGGGGAGTGGCTGCTGGCGGAGCGCTGTGCCAC 360
QY 395 GTGTGGATCTCTTCGAGCGCGAGCGCTGT-----CTGTG 429
Db 361 GTGTGGATCTCTTCGAGCGCGAGCGCTGTGCCAGTGTGATCTCTTCACGCGTGTG 420
QY 430 CTGCCCCCGCGCTCGGGAAAGTGGCGCCATGCCCTGGCGCGCGACGGGGCAATCAC 489
Db 421 CTGCCCCCGCGCTCGGGAAAGTGGCGCCATGCCCTGGCGCGCGACGGGGCAATCAC 480
QY 490 ACGGCACTGACAGACAGCTGCGGACCGGACCGCGCGCTGCTCATGATCGGCT 549
Db 481 ACGGCACTGACAGACAGCTGCGGACCGGACCGCGCGCTGCTCATGATCGGCT 540
QY 550 CGCCCGGCTGCGCTGCGGCTCATCGCCCTCGCGCGCTGCTCTTGGCGGGGCGAGT 609
Db 541 CGCCCGGCTGCGCTGCGGCTCATCGCCCTCGCGCGCTGCTCTTGGCGGGGCGAGT 600
QY 610 GTGCGAGCTCGGCTCCAGCGCTGCCAGGTGAGCGGAAACCTCTATGCGGCTTCTC 669
Db 601 GTGCGAGCTCGGCTCCAGCGCTGCCAGGTGAGCGGAAACCTCTATGCGGCTTCTC 660
QY 670 CACCGCGGCGCTTCCAGCTGCGCTGGGCTGGTGGCTGCTGCTACCGGAGATCTA 729
Db 661 CACCGCGGCGCTTCCAGCTGCGCTGGGCTGGTGGCTGCTGCTACCGGAGATCTA 720
QY 730 CGAGCGCGCAAGTTTCGTTTTCGCGCGCGCGGAGAGCTGTGCTGCGGCGCAC 789
Db 721 CGAGCGCGCAAGTTTCGTTTTCGCGCGCGCGGAGAGCTGTGCTGCGGCGCAC 780
QY 790 CATGCAAGTGAAGTTCAGAGTAAAGGACCACT 823
Db 781 CATGCAAGTGAAGTTCGAGTGAAGGACCACT 814

RESULT 14

ABKS0434

ID ABKS0434 standard; DNA; 1074 BP.

XX AC ABKS0434;

XX DT 30-JUL-2002 (first entry)

XX DE Human 5-hydroxytryptamine receptor 5A (HTR5A) DNA coding region.

XX KW Human; 5-hydroxytryptamine receptor 5A; HTR5A; serotonin; gene; ds;

XX KW neuroprotective; neurological disease; depression; epilepsy;

XX KW gene therapy; single nucleotide polymorphism; haplotype pair;

XX KW Chromosome 7q36.1.

XX OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1074

XX FT /*tag= a

XX FT /product= "Human HTR5A"

XX FT replace(12,T)

XX FT /*tag= b

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(43,T)

XX FT /*tag= c

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(300,C)

XX FT /*tag= d

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(358,C)

XX FT /*tag= e

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(459,G)

XX FT /*tag= f

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(469,A)

XX FT /*tag= g

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(789,T)

XX FT /*tag= h

XX FT /standard_name= "Single nucleotide polymorphism"

XX FT replace(849,T)

XX FT /*tag= i

XX FT /standard_name= "Single nucleotide polymorphism"

XX WO200222887-A1.

XX PD 21-MAR-2002.

XX PF 17-SEP-2001; 2001WO-US029210.

XX PR 15-SEP-2000; 2000US-0233051P.

XX PA (GENA-) GENAISSANCE PHARM INC.

XX PI Kazemi A, Koshy B, Sanchis A, Tirrell C;

XX DR WPI; 2002-393978/42.

XX DR P-PSDB; AAU79252.

XX Novel genetic variants of 5-Hydroxytryptamine (Serotonin) Receptor 5A

XX PT isogenes, useful for improving efficiency and reliability in drug

XX PT development for treating neurological diseases.

XX PS Claim 27; Fig 3; 134pp; English.

XX The invention relates to single nucleotide polymorphisms in the gene

XX CC encoding human 5-hydroxytryptamine (serotonin) receptor 5A (HTR5A). A

XX CC method for haplotyping the HTR5A gene in an individual comprises

XX CC identifying the nucleotide at one or more polymorphic sites and

CC avidity for a particular GPCR. (1) can be used as GPCR modulators and in
CC gene therapy. The antigenic peptides for GPCRs are useful in detecting an
CC antibody against a particular GPCR, and in the production of specific
CC antibodies. The peptides and antibodies are also useful for detecting the
CC presence or absence of corresponding GPCRs. The antigenic peptides for
CC GPCRs and antibodies are useful for diagnosing and designing drugs for
CC treating immune-related diseases, growth-related diseases, cell
CC regeneration-related disease, immunological-related cell proliferative
CC diseases, or autoimmune diseases, e.g. AIDS, Alzheimer's disease,
CC atherosclerosis, bacterial, fungal, protozoan or viral infections,
CC osteoarthritis, osteoporosis, cancer, cardiomyopathy, chronic and acute
CC inflammation, allergies, Crohn's disease, diabetes, graft versus host
CC disease, Parkinson's disease, multiple sclerosis, pain, psoriasis,
CC anxiety, depression, schizophrenia, dementia, mental retardation, memory
CC loss, epilepsy, asthma, tuberculosis, obesity, nausea, hypertension, or
CC hypotension, renal disorders, rheumatoid arthritis, trauma, ulcers, or
CC any other disorder in which GPCRs are involved. The antibodies may be
CC used in immunoassays and immunodiagnosis. ABZ42523 to ABZ42869 encode
CC GPCR proteins given in ABP81675 to ABP82018, which are used in the
CC exemplification of the present invention
XX
SQ

Sequence 1074 BP; 182 A; 366 C; 284 G; 242 T; 0 U; 0 Other;

Query Match 43.2%; Score 497.6; DB 7; Length 1074;

Best Local Similarity 70.6%; Pred. No. 3.6e-98;

Matches 729; Conservative 0; Mismatches 279; Indels 24; Gaps 4;

Qy	113	ACCCGAGCGGCGCTCTGCGCGGCGGAGGCGCGCTCTCTGCTGCTTCACGGTCTGTG	172
Db	67	AGCCTCGGCAAGACGACCTGCGCGCCCTGCTCTCGGCTCTCGGAGTGTCT	126
Qy	173	GTGCTGACGCTGTAGTGTCTGATCGTSCCACTTCTCTGTGAACTGCTGTTCCG	232
Db	127	ATTCTACCTTGTGGGCTTCTGCTGGCGGACGTTGGCTGGAACCTGCTGTGTG	186
Qy	233	GTACCAATCCCGGGTCCGTGCTTCCACCGCTGCGGATAACTTGGTGGCTTCGACG	292
Db	187	GCGACCATCTCCGTGTACGACCTTCCACCGCTGCCCCACACCTGGTGGATCCATG	246
Qy	293	GCGCTCTCGGACGAACTAGTGGGAGCGGTGGCGATGCCACGAGCCTGGCGAGTGTG	352
Db	247	GCGCTCTCGGATGTCCTGCTGGCGCGCTGGTCTATGCGGCTTGAGCCTGGTGCATG	306
Qy	353	TCGACCGGGGCGCTGCGGCTGCTGGCGGAGCCTGTGSCACGCTGATCTCTCGAC	412
Db	307	T---CCGGGCGCGCTGGCAGTAGTTCGGAGGCTGTGCCAGCTTTGGATCCGCTGC	363
Qy	413	GCGGAGCCTGTGTGTGCTGCCCGCGGCTCGGGAACTGGCGGCAATCGCCCTGGGC	472
Db	364	G-----TGCCTTGTGCGACGGCGAGCATCTGGAACGTGACGGCCATAGCCCTGGAC	414
Qy	473	GCGACCGGGGCGCATCACACGCACTGACAGACACGCTGGCGACCCCGAGCGCGCTCG	532
Db	415	CGCTACTGGTCCATCACGCGGCACATGGAATACACGCTCCGCAAGTGGCTCTCC	474
Qy	533	TTGCTCATGATCGCGCTGCCCGGGTGCCTGCGGCTCATCGCCCTCGCGCGCTGCTC	592
Db	475	AACGTCAATGATCGGCTCACTGGGCACTCTCCGCTGTCACTCTTGGCCCCCGCTGCTT	534
Qy	593	TTTGGCCGGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCGGGAAACC	652
Db	535	TTTGGCTGGGGAGAGAGCTACTCTGAGGCGAGCGAGGAGTSCCAGGTAAAGCCGAGCCT	594
Qy	653	TCCTATGCGGCTTCTCCACCGCGGCGCTTCCACCTGCGGCTTGGGTGGTGGCGTTT	712
Db	595	TCCTACGCGGTGTTCTCCACCGTAGGGCGCTTCTACCTGCGGCTCTGTGTGTCTCTC	654
Qy	713	GTCTACCGGAAGATCTACGAGCGGCGCAAGTTTCGTTTCGGCGCGCGCGGAGAGCTGTG	772
Db	655	GTGTACTGGGAAGATCTACAGGCTGCGCAAGTTCGCGTGGGCTCCAGGAAGCAATAGC	714
Qy	773	CTGCCGTTGCCGGCCACCATGCAAGTGTAGGTTCAAGGTAAAGGAAGCACTGTAGGCT	832

Db	715	GT-----CTCACCCATATCGAAGCTGTGTGAGGTGAAGGACTCTGCCAAACAGCCC	765
Qy	833	GAACTGGTGTTCAGCGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCC	892
Db	766	CAGATGGTGTTCAGGTCC---GCCAGCCACCGTCACTTCCAGCCAGAAAGGGGACAG	822
Qy	893	TGGCGGAGCAGAAAGAGAGCGAGCCATGATGTGTGGGAAATCTGATTTGGCGTGT	952
Db	823	TGGCGGAGCAGAAAGAGAGCGGCGCCCTCATGTGTGGCATCCTCATTTGGCGTGTTC	882
Qy	953	GTGCTGTGTGGATCCCTTCTTCTTCCAGCGAACTCATCAGCCCACTGTGTGCTGTGAGC	1012
Db	883	GTGCTGTGTGGATCCCTTCTTCTTCCAGCGAGCTCATCAGTCCCTCTGTCTCTGTGAC	942
Qy	1013	CTGCCCCCATCTGGAAGCATATTTCTGTGGCTTGGCTACTCCAATTTCTTCTTCAAC	1072
Db	943	ATCCCCCCCATCTGGAAGCATCTTCTTGTGGCTTGGCTACTCCAATCTTCTTTAAC	1002
Qy	1073	CCCCTGATTTACACAGCTTTTAAACAAGAACTACAACTGCTTCAAGAGCCTCTTACT	1132
Db	1003	CCCCTGATCTATAGGCTTTTCAACAAGAACTACAAAGCGCTTCAAGAACTCTTTTCT	1062
Qy	1133	AAGCAGAGATGA	1144
Db	1063	AGGCAACACTGA	1074

Search completed: September 7, 2004, 15:10:54
Job time : 507 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 15:11:05 ; Search time 2862 Seconds
(without alignments)
12019.997 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence: 1 cgccatgagcgccgtagcc.....aagcagatgacacacaggg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 segs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estcov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	496	43.1	1074	29	AY402561	Homo sapi
2	495	43.0	1074	29	AY402562	Pan trogl
c 3	477.6	41.5	640	28	BZ601879	WHAAAS5TF
4	426.2	37.0	1074	29	AY402563	Mus muscu

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

5	284.2	24.7	1099	12	BM805859
6	280.4	24.3	906	12	BI489745
7	265.4	23.0	539	13	BY280849
8	255.6	22.2	1027	29	CNS05MTV
c 9	255.2	22.2	893	12	BI603546
c 10	255.2	22.2	2683	11	BC031272
c 11	247.6	21.5	583	12	BI538349
c 12	239.6	20.8	472	12	BG715343
c 13	200.4	17.4	929	12	BI411238
c 14	198.4	17.2	660	10	BB628489
c 15	173.8	15.1	655	12	BM861710
c 16	149.4	13.0	316	14	CE600820
c 17	148.2	12.9	744	29	CE490868
c 18	144.2	12.5	646	12	BM861408
c 19	140.8	12.2	730	12	BI820920
c 20	120.2	10.4	538	9	AL922333
c 21	112.6	9.8	874	14	CB565883
c 22	107	9.3	697	10	BB653529
c 23	96.6	8.4	1213	29	AY402238
c 24	93.6	8.1	988	29	CNS018EH
c 25	93.6	8.1	1269	29	AY402237
c 26	93	8.1	503	14	CB784054
c 27	93	8.1	745	12	EG827135
c 28	93	8.1	780	12	BI195997
c 29	93	8.1	914	13	BQ879192
c 30	93	8.1	929	13	BQ719305
c 31	93	8.1	1332	29	AY418851
c 32	91.4	7.9	519	28	AQ897350
c 33	91.4	7.9	1335	29	AY418853
c 34	91.4	7.9	1201	9	AL530418
c 35	90	7.8	1058	12	BM926157
c 36	88.8	7.7	1221	29	AY409288
c 37	88.8	7.7	1323	29	AY417222
c 38	86.6	7.5	1201	9	AL544609
c 39	86.4	7.5	1221	29	AY409287
c 40	84.4	7.3	1764	11	BC018330
c 41	84.4	7.3	2611	11	AK004891
c 42	84.4	7.3	3050	11	AK043877
c 43	83.8	7.3	1072	29	AY398880
c 44	83	7.2	1401	29	AY399426
c 45	82.4	7.2	1125	29	AY418220

ALIGNMENTS

RESULT 1	AY402561	1074 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY402561	Homo sapiens HTR5A gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.			
DEFINITION	AY402561				
ACCESSION	AY402561				
VERSION	AY402561.1	GI:39758547			
KEYWORDS	GSS.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	1	(bases 1 to 1074)			
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2	(bases 1 to 1074)			
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1074
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
<1..>1074
/gene="HTR5A"
/locus_tag="HCML265"
ORIGIN
Query Match 43.1%; Score 496; DB 29; Length 1074;
Best Local Similarity 70.5%; Pred. No. 2.2e-80;
Matches 728; Conservative 0; Mismatches 280; Indels 24; Gaps 4;
QY 113 ACCCCGAGCGCGCGCTCTGCGCGGCGCGAGGCGCGCCCTTCTCTGTCTTCTCAGGTCCTG 172
Db 67 AGCCTCGGCAAGACGACCTGCGCCCGAGCTGCGCCCTGCTCTCGGTCTTCGGAGTGCTT 126
QY 173 GTGGTACGCTGTAGTGTCTGATGCTGCTGCGTGCACCTTTCTCTGTGGAACCTGTGTTCCG 232
Db 127 ATTCTCACCTTGTCTGGGCTTTCTGTGGCGCGACGTTGCGCTTGGAACTGTGTGTGCTG 186
QY 233 GTACCATCCGCGGCTCGTGTCTTCCACCGGTGCGCATAACTTGTGTGGCTCGACG 292
Db 187 GCGACCATCTTCGTGTACGACCTTCCACCGGTGCGCCACAACTGTGTGGCATCCATG 246
QY 293 GCGCTCTCGGACAACTAGTGGGAGCGCTGGCGATGCCACCGACCTTGGCGAGTGAGCTG 352
Db 247 GCGCTCTCGATGCTCTGTGTGGCGCGCTGGTCTATGCGCGTGAAGCTGTGTGACGAGCTG 306
QY 353 TCGACCGGCGGACGTCGCTGTGTGGCGGAGCTGTGCCACGCTGTGTGATCTCTCTTCGAC 412
Db 307 T---CGGGCGCGCTGGCAGCTAGGTTCGGAGGCTGTGCCAGCTTTGGATCGCGTGGCAG 363
QY 413 GCGGAGGCTGTCTGTGTGCTGCGCGCGCTCGGGAACGTGGCGGCATCGCCCTGGGC 472
Db 364 G-----TGTTTGTCTGACGGCGAGCATCTGGAACGTGACGGCCATAGCCCTTGGAC 414
QY 473 CGGACGGGSCCATCACGCGACCTCGACACACGCTGGCGACCGCGCGCGCTCG 532
Db 415 CGCTACTGTGTCCATCAGCGCCACATGGAATACACGCTCGCACCCGCAAGTCGCTCC 474
QY 533 TTGCTCATGATCGCGTGTGCGCGGCTGCGTGGCGCTATATCGCCCTGCGCGCTGCTC 592
Db 475 AAGCTCATGATCGCGTCACTCGGCACTCTCGCTGTCTATCTCTCTGCGCCCGCTGCTT 534
QY 593 TTTGGCGGCGGAGGTGTGCGAGCTCGGCTCCAGCGCTCGCAGGTGAGCGGGGAACCC 652
Db 535 TTTGGCTGGGAGAGACGTACTCTGAGGCGAGGAGGTCCAGGTAGACCGCGAGCCT 594
QY 653 TCTATTCGCGCTTCTCCACCGCGGCGCTTCCACCTGCGGCTTGGCGGTGTGCGGTTT 712
Db 595 TCCTACGCGGTCTTCCACCGTAGCGCTTCTACTGCGGCTCTGTGTGTGCTCTTC 654
QY 713 GTCTACGGAAGATCTACGAGGCGGCAAGTTTCTTTCGCGCGCGCGGAGAGCTGTG 772
Db 655 GTGTACTGGAAGATCTACAAAGGTGCAAGTTTCGCGTGGGCTCCAGGAAGCAATAGC 714
QY 773 CTGCGGTTCGCGGCCACCATGCAAGTGTGAGTCCAAAGTAAAGAACACCTGATGAGGT 832
Db 715 GT-----CTCACCATATCGAAGCTGTGAGGTGTGAGGCTGTGCAACAGGCC 765
QY 833 GAAGTGTGTTCACGACATTCGAAAGCAACCGGTGTCTTCACAGGTGAGCGGGACTCC 892
Db 766 CAGATGTTTTCACGGTCC---GCCACGCCACCGTCACCTTCCAGCGCAGAAAGGGACACG 822
QY 893 TCGCGGAGCAGAGGAGCGGACGACCATGATGTTGGGAATTCGTATTTGCGGTGTTT 952
Db 823 TGGCGGAGCAGAAAGAGCAGCGGCGCGCCCTCATGTGTGGGCATCTCTCATTTGCGGTTC 882

QY 953 GTGCTGTGTGATCCCTTCTTCTGACGGAACCTCATCAGCCCACTCTGTGCTCGACG 1012
Db 883 GTGCTGTGTGATCCCTTCTTCTCAGCGAGCTCATCAGTCCCTCTGCTCTGTGAC 942
QY 1013 CTGCCCCCATCTGGAAGACATATTTCTGTGGCTGTGCTACTCCAAATCTTTTCTCAAC 1072
Db 943 ATCCCGCCCATCTGGAAGACATCTTCTGTGGCTGTGCTACTCCAACTCTTCTTTAAC 1002
QY 1073 CCCCTGATTACACAGCTTTTAAACAGAACTACAACAATGCCTTCAAGAGCCTCTTACT 1132
Db 1003 CCCCTGATCTATACGGCTTTTCAACAAGAACTACAACAGCGCTTCAAGAACTCTTTTCT 1062
QY 1133 AAGCAGAGATGA 1144
Db 1063 AGCAACACTGA 1074
AV402562 1074 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes HTR5A gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402562
VERSION AY402562.1 GI:39758548
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1074)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1074)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sainsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source Location/Qualifiers
1..1074
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>1074
/gene="HTR5A"
/locus_tag="HCML265"
ORIGIN
Query Match 43.0%; Score 495; DB 29; Length 1074;
Best Local Similarity 70.4%; Pred. No. 3.3e-80;
Matches 727; Conservative 0; Mismatches 281; Indels 24; Gaps 4;
QY 113 ACCCCGAGCGCGCGCTCTGCGCGGCGCGAGGCGCGCCCTTCTCTTCTTCTCAGGTCCTG 172
Db 67 AGCCTCGGAAAGACGACCTGCGCCCGAGCTGCGCCCTGCTCTCGGTCTTCGGAGTGCTT 126
QY 173 GTGGTACGCTGTAGTGTCTGATGCTGCTGCGACATTTCTCTGTGGAACCTGTGTGTTCCG 232
Db 127 ATTCTCACCTTGTGTGGGCTTTCTGTGTGGCGGACGTTGCGCTTGGAACTGTGTGTGCTG 186
QY 233 GTACCATCCGCGGCTCGTGTCTTCCACCGGTGCGCATAACTTGTGTGGCTCGACG 292

Db 187 GCGACCATCTCCGTGTACGACCTTCCACCGGTGCCCCCAACAACTGGTGATCCATG 246

Qy 293 GCGTCTGGAGCACTAGTGGCAGGCTGGGATGCCACCGAGCTGGCAGTGGAGCTG 352

Db 247 GCGATCTGGATGCTGTGGCCGCGCTGGTATGCCGCTGAGCTGGTGACAGCTG 306

Qy 353 TCGACCGGCGACGTGGCTGTGGGCGCGAGCCCTGTGCGACGTGTGGATCTCTTCGAC 412

Db 307 T---CGGCGCGCTGGCAGCTAGGTGGAGGCTGTGCCAGCTTTGGATCTCGTGGAC 363

Qy 413 GCGGAGCTGTCTGTGCTGCCCGCGCGCTGGGAGAGCTGGGCGCATGCGCTGGGC 472

Db 364 G-----TGCTTTGTGTCAGCGCCAGCATCTGGAACGTGACGCCATAGCCCTGGAC 414

Qy 473 GCGGAGCGGCGCATCACAGCGCACCTGCGACACAGCTGCGCACCGCAGCGCGCCCTCG 532

Db 415 CGCTACTGTGCATCAGCGGCGCATGGAATACAGCTCCGACCGCAAGTGGCTGCC 474

Qy 533 TTGCTCATGATCGCGCTCGCCCGGCTGGCGGCTCATCGCCCTGCGCGCTGCTGCT 592

Db 475 AACGTGATGATCGCGCTCACCTGGGCACTCTCCGCTGTCTCTCTGCGCCGCTGCTT 534

Qy 593 TTGCGCGGCGGAGTGGAGCGCTCGCTCGCTCCAGCGCTGCGAGTGAGCGGGAACCC 652

Db 535 TTGCTGGGAGAGACGCTACTCTGAGGCGACGCGAGAGTGCAGGTAAGCGCGAGGCT 594

Qy 653 TCCTATGCGCGCTCTCCACCGCGCGCTTCCACCTGCGCTTGGGTGGTGGCGGTTT 712

Db 595 TCCTAGCGCGTCTCCACCGTGGCGCTTCTACCTGCGCTGTGTGGTGGTCTCTTC 654

Qy 713 GTCTACCGGAAGATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGGAGAGCTGTG 772

Db 655 GTGTAAGTGAAGATCTACAAGGCTGCCAAGTTCGCGGTCCAGGTAAGCAATAGC 714

Qy 773 CTGCGGTCGCGGCGCACATGACGAGTGAAGTCCAAAGTAAAGGAAGCACTGATGAGCT 832

Db 715 GT-----CTCACCCATATCCGAAGCTGTGGAGTGAAGGANTCTGCCAAACAGGCC 765

Qy 833 GAAGTGGTGTTCACGGCACATTTGAAAGCAAGGTGTCTTCCAGTGGCGGGACCTCC 892

Db 766 CAGATGGTTCACGGTCC---GCCACGCCACCGTCACTTCCAGACAGAGGGACACG 822

Qy 893 TGGCGGAGCAGAGAGGAGGCGAGCAGCCATGATGTTGGGAATTCGATTGGCGTGTTC 952

Db 823 TGGCGGAGCAGAGAGGAGCAGCGGCGCCCTCATGGTGGGATCCTCATTTGGTGTTC 882

Qy 953 GTGCTGTGCTGATCCCTTCTCTGACGGAATCATCAGCCGACTCTGTGCGCTGACG 1012

Db 883 GTGCTGTGCTGATCCCTTCTCTCACCGAGCTCATCAGTCCCTCTGCTCTCTGAC 942

Qy 1013 CTGCCCCCATCTGGAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTCTTCAAC 1072

Db 943 ATCCCCACCATCTGGAAGGATCTTCTGTGGCTTGGCTACTCCAACTCTTCTTAAAC 1002

Qy 1073 CCCCTGATTTACAGCTTTTAAAGAAGTACAAATGCTTCAAGAGGCTCTTTTACT 1132

Db 1003 CCCCTGATCTATAGCGCTTTCAAGAAGTACAAAGTGTGCTTCAAGAACTTATTTCT 1062

Qy 1133 AAGCAGATGA 1144

Db 1063 AGGCAACACTGA 1074

RESULT 3

LOCUS BZ601879/c 640 bp DNA linear GSS 08-JUN-2003

DEFINITION WHAA55TF Human MCF7 breast cancer cell line library (MCF7.1) Homo sapiens genomic clone MCF7_1-1113, genomic survey sequence.

ACCESSION BZ601879

VERSION BZ601879.1 GI:31510341

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 640)

Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q., Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P., Gray, J.W. and Collins, C.

End-sequence profiling: Sequence-based analysis of aberrant genomes

Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)

22709111

12788976

Contact: Volik SV

Colin Collins' lab

UCSF Comprehensive Cancer Center

UCSF Box 0808, San Francisco, CA 94143-0808, USA

Tel: 415 502 7066

Fax: 415 502 5665

Email: svolik@cc.ucsf.edu

This clone is available from Amplicon Express

http://www.genomex.com

Class: BAC ends.

Location/Qualifiers

1..640

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clone="MCF7_1-1113"

/sex="female"

/clone_lib="Human MCF7 breast cancer cell line library (MCF7.1)"

/note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 41.5%; Score 477.6; DB 28; Length 640;

Best Local Similarity 92.2%; Pred. No. 4.3e-77;

Matches 535; Conservative 0; Mismatches 19; Indels 26; Gaps 2;

Qy 270 GCGATACTTGGTGGCTCGACGGCCGCTCGGACGAACTAGTGGCAGCGTGGCGATGC 329

Db 640 GCGATACTTGGTGGCTCGACGGCCGCTCGGACGAACTAGTGGCAGCGTGGCGATGC 581

Qy 330 CACGAG--CTGGGCGAGTGAAGTGTGACCGGCGGACGCTCGGCTGCGCGGAGCCTG 388

Db 580 CACGAGCCCTGGCGAGTGAAGTGTGACCGGCGGACGCTCGGCTGCGCGGAGCCTG 521

Qy 389 TGCCACGTGTGATCTCCTTCGACGCGGAGCCTGT----- 424

Db 520 TGCCACGTGTGATCTCCTTCGACGCGGAGCCTGTGCGGATCTCCTTCCAG 461

Qy 425 -CTGTGCTGCCCGCGGCTCGGGAACGTGGGCGCATCGCCTGTGGCGCGAGCGGC 483

Db 460 GCTGTGCTGCCCGCGGCTCGGGAACGTGGGCGCATCGCCTGTGGCGCGAGCGGC 401

Qy 484 CATCACGCGCACCTGCGAGCACGCTGCGACCGGAGCGCGCTGCTGCTCATGAT 543

Db 400 CATCACCGCACCTGCGAGCACGCTGCGACCGGAGCGCGCTGCTGCTCATGAT 341

Qy 544 CGCGCTCGCCCGGCTGCGGCTCATCGCCTCGCGCGCTGCTTGTGGCGGG 603

Db 340 CGCGCTCACCGGCTGCTGCTCGGCGCTCATCGCCTCGCGCGCTGCTTGTGGCGGG 281

Qy 604 CGAGGTGTGCAAGCTCGCTCGCTCCAGCGCTGCCAGTGAGCGGGAACCTCTATGCGCG 663

Db 280 CGAGGTGTGCAAGCTCGCTCGCTCCAGCGCTGCCAGTGAGCGGGAACCTCTATGCGCG 221

Qy 664 CTTCTCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGCTTGTCTACCGGAA 723

Db 220 CTTCTCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGCTTGTCTACCGGAA 161

Qy 724 GATCTACGAGCGGCCCAAGTTTCCTTTTCGCGCGCGCGGAGAGCTGTGCTGCGCTTGC 783

```

Db      160 GATCTACGAGCGGCCAAGTTTCGTTTCGGCCGCCCGGAGAGCTGTGTCGCCGTTGCC 101
QY      784 GCCACCATCGAGTTCAGTCCAGGTAAGGAAGACACT 823
Db      100 GGCCACCATCGAGTTCAGGTTGGGCTGAGGAACGTTGCT 61

RESULT 4
AY402563      1074 bp      DNA      linear      GSS 15-DEC-2003
LOCUS      Mus musculus HTR5A gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION      genomic survey sequence.
ACCESSION      AY402563
VERSION      AY402563.1 GI:39758549
KEYWORDS      GSS.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1074)
Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE      Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL      Science 302 (5652), 1960-1963 (2003)
PUBMED      14671302
REFERENCE      2 (bases 1 to 1074)
AUTHORS      Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.
TITLE      Direct Submission
JOURNAL      Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
COMMENT      Rockville, MD 20850, USA
      This sequence was made by sequencing genomic exons and ordering
      them based on alignment.
FEATURES
source      Location/Qualifiers
1..1074
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1074
/gene="HTR5A"
/locus_tag="HOM1265"
ORIGIN

Query Match      37.0%; Score 426.2; DB 29; Length 1074;
Best Local Similarity 67.4%; Pred. No. 1.2e-67;
Matches 685; Conservative 0; Mismatches 308; Indels 24; Gaps 5;

QY      128 GTCTCGCGGCGGCGGCGGCGGCTTCTCTGTCTTTCACGGTCTCTGTGTGTGAGCGTGTCTA 187
Db      82 GTCTCGCGGCGGCGGCGGCTTCTCTGTCTTTCACGGTCTCTGTGTGTGAGCGTGTCTTGTG 141
QY      188 GTCTCGCTGATCGCTGCACCTTTCCTGTGGAACCTGCTGCTTCCGGTCCACATCCCGCGG 247
Db      142 GCGCTTCTAGCTCGGCGCACATTCATCTTGGAACCTGCTGTGTGTGCTACCATCTCTCAAG 201
QY      248 GTCGCTGCTTCCACCGCGTCCGCATAACTTGGTGGCCTTCGACGCGCGTCTCGGACGAA 307
Db      202 GTACGCACCTTCCACCGAGTACCACACAACTGTAGCTTCCATGGCCATCTCGGATGTG 261
QY      308 CTAGTGGCAGCTGGGATGCCACCGAGCTGGCAGTGTGTCTGACGCGGCGGACGCT 367
Db      262 CTAGTGGCTGTGCTGTATGGCCACTGAGCTGTGTATCATGAGCTGT---CTGGGCGCGC 318
QY      368 CGCTGTGCTGGGCGGAGCGTGTGCCAGTGTGTGATCTCTTCAGCGCGGAGCGTGTGTG 427
Db      319 TGGCAGCTGGGCGGCGGCTCTATGCCAGCTGTGGATCGCATGTGACG-----TGCTC 369
QY      428 TGCTGCCCCCGCGGCTTCGGGAACGTTGGCGGCCATGCGCCCTGGGCGCGGAGCGGCGCATC 487

```

```

Db      370 TGCTGTACTGCCAGCATCTGGAATGTGACAGCAATAGCTCTGTGACCGCTACTGGTCAATC 429
QY      488 ACACGGCAGCTGCAGACACAGCTGGCAGCCGCGGCGGCGCTGTGCTCATGATCGCG 547
Db      430 ACGCGCCACCTGGAGTACACACTCCGTACCCGCAAGCGTGTCTCCAATGTGATGATCCTG 489
QY      548 CTCGCCCGGCTGCGGTGCGGCTCATGCGCCCTTCGCGCCGCTGTCTTTTGGCGGGCGGAG 607
Db      490 CTCACCTGGGCACTCTCCACTGTCACTCTCTGGCTCCACTGCTATTGTGGTGGGAGAG 549
QY      608 GTGTGGAGCCTCGGCTCCAGCGCTGCAGGTGAGCCGGGAACCTCTCTATGCCGCTTC 667
Db      550 ACTTATTCTAGCCCAAGTGCAGGAATGCCAAGTCAGTCGCGAGCTTCTTACACCGTGTTC 609
QY      668 TCCACCCGCGGCGCTTCCACCTGCGCTTGGCTGTGCGCTGCTCTTCTTACCGGAAGATC 727
Db      610 TCACCGTGGGTGCGCTTCTACCTGCGCGTGTGGTGTGCTCTTTGTGCTAGGAAAT 669
QY      728 TAGGAGCGGCGCAAGTTTCGTTTTCGCGCGCGCGGAGAGCTGTGCTGCGGTGCGCGGC 787
Db      670 TACAGGCGCGCAAAATTCGCAATGGGCTCCAG--GAAGACCAACAGCGTCTCCCGCTAC 727
QY      788 ACATGCGAGTGAAGTCCAAAGTAAGGAAGCACCTGATGAGGCTGAAGTGTGTTCACG 847
Db      728 CCGAAGCTGTGG-----AGGTGAAGAATGCTTACACAACATCCCGAGATGGTGTTCACG 780
QY      848 GCACATTGCAAAAGCAACGCTGCTCTTCAGGTGAGCGGGGACTCTCCGCGGAGCAGAAG 907
Db      781 GTCC---GCCATGCCACCGTCACTTCCAGAGAAGGGATACGTGGAGGAGCAGAAG 837
QY      908 GAGAGCGAGCAGCAGCATGATGTGGAAATCTGATTGGCGTGTGTGCTGTGCTGTGATC 967
Db      838 GAGCAAGGCGCAGCCCTCATGTTGGGCATCCTCATCGAGTGTGTGTGCTGTCTGTGTT 897
QY      968 CCCTTCTTCTGACGGAACATCATCAGCCACTCTGTGCTGTGAGCTGCCCCCATCTGG 1027
Db      898 CCTTTCTCGTCACAGAGCTCATCGTCCCTGTGTCTCTGGGAGCTCCCTGCGCATCTGG 957
QY      1028 AAGAAGCATATTTCTGTGGCTTGCTACTCCTCAATCTTCTTCAACCCCTGATTACACA 1087
Db      958 AAGAGCATCTTCTGTGGTTGGGCTATTCTAATCTCTTCAACCCATCATCTACACA 1017
QY      1088 GCTTTTAAACAAGAACTACAAATGCTTCAAGAGCGCTCTTTACTAAGCAGAGATGA 1144
Db      1018 GCATTCAACAGGAGCTACAGCAGTGTCTTCAAGGTCTTCTTCTCCAAGCAACAATGA 1074

RESULT 5
BM805859
LOCUS      BM805859
DEFINITION      AGENCOURT_6513650 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5731557
S', mRNA sequence.
ACCESSION      BM805859
VERSION      BM805859.1 GI:19122682
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 1099)
AUTHORS      NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL      Unpublished (1999)
COMMENT      Contact: Robert Strausberg, Ph.D.
      Email: cgabbs-f@mail.nih.gov
      Tissue Procurement: Invitrogen
      cDNA Library Preparation: Life Technologies, Inc.
      cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
      DNA Sequencing by: Agencourt Bioscience Corporation
      Clone distribution: MGC clone distribution information can be
      found through the I.M.A.G.E. Consortium/LNL at:
      http://image.lnl.gov

```

Plate: LLAM12732 row: d column: 22
High quality sequence start: 7
High quality sequence stop: 616.
Location/Qualifiers
1. 1099
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5731557"
/tissue_type="hippocampus"
/lab_host="DH10B"
/clone_lib="NIH_MGC 124"
/note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV (destroyed); Site 2: NotI; RNA source male hippocampus, age 27. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 0.9-4 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 012."

ORIGIN

Query Match 24.7%; Score 284.2; DB 12; Length 1099;
Best Local Similarity 71.2%; Pred. No. 1.1e-41;
Matches 408; Conservative 0; Mismatches 153; Indels 12; Gaps 2;
Qy 572 ATCCCTCGGCGCGCTCTTTGGCGGGGAGGTGTGGAGCTCGGCTCCAGGC 631
Db 18 ATCTCTGCGCGCGCTCTTTGGCTGGGAGAGACGTACTCTGAGGCGAGCGAG 77
Qy 632 TGCCAGGTGAGCGGAGACCTCTATGCCGCTTCTCCACCGCGGGCGCTTCCACCTG 691
Db 78 TGCCAGGTGAGCGGAGACCTCTATGCCGCTTCTCCACCGGTAGGCGCTTCTACCTG 137
Qy 692 CCGCTTGCGCGTGTGCGCTTGTCTACCGAAGATCTACGAGGCGGCGCAAGTTGTTTC 751
Db 138 CCGCTTGCGGTGTGCTCTCGTGTACTGGAAGATCTCAAGGTGCGCAAGTTGCGGTG 197
Qy 752 GCGCGCGCGGAGAGCTGTCTGCGGTGCGGCGCACCATGCGAGGTGAGGTCCAGGTA 811
Db 198 GCGTCCAGGAAGACCAATAGCGT-----CTCACCCATATCCGAAGCTGTGGAGGTG 248
Qy 812 AAGGAAGCACTGATGAGGCTGAAGTGTGTTACGGGCACATTCGAAAGCAAGCTGTGCC 871
Db 249 AAGGACTGTGCAACAGCCCGACAGTGTGTTCAGGTGCG---GCCAGCCACCGTCAAC 305
Qy 872 TTCCAGGTGAGCGGGAATCTCTGCGGAGCAGAAAGAGAGCGAGCGCATGATGTTG 931
Db 306 TTCCAGCCAGAAAGGACACGTGGCGGAGCAGAAAGAGAGCAGCGGCGCGCTCATGGTG 365
Qy 932 GGAATTCGATGGGTTGT 991
Db 366 GGCATCTCATGGCGT 425
Qy 992 AGCCCACTGTGCTGCGAGCTGCGCCCGCATCTGGAAGAGATATTTCTGTGGCTTGGC 1051
Db 426 AGTCCCTCTGCTCTGTGACATCCCGGCATCTGGAAGAGATCTTCTGTGGCTTGGC 485
Qy 1052 TACTCCAATCTTTCTTCAACCCCTGATTTTACAGCTTTTAAAGAACTTAAACAAAT 1111
Db 486 TACTCCAATCTTTCTTAAACCCCTGATCTATACGGCTTTCAACAAAGAACTTAAACAGC 545
Qy 1112 GCGTTCAAGAGCTCTTTACTAAGCAGATGA 1144
Db 546 GCGTTCAAGAACTTTCTTTCTAGGCAACACTGA 578

RESULT 6

BI489745
LOCUS
DEFINITION
603032245F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5173354 5',
mRNA sequence.
ACCESSION
BI489745

VERSION
KEYWORDS
SOURCE
ORGANISM

BI489745.1 GI:15328973
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgs.nci.nih.gov/.
1 (bases 1 to 906)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11431 row: j column: 11
High quality sequence stop: 826.

FEATURES

source

Location/Qualifiers

1. 906
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5173354"
/lab_host="DH10B"
/clone_lib="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 24.3%; Score 280.4; DB 12; Length 906;
Best Local Similarity 71.1%; Pred. No. 5e-41;
Matches 431; Conservative 0; Mismatches 161; Indels 14; Gaps 4;
Qy 540 TGATCGCGCTCGCGGTCGCGGCTCATCGCCTCGCGCGCTGCTCTTTGGCC 599
Db 1 TGATCGCGCTCACCTGGGCACTCTCCGCTGTCTCTGCGCCCGCTGCT-TTTGGCT 59
Qy 600 GGGGCGAGGTGTGCACGCTCGGCTTCCAGCGCTGCCAGGTGAGCGGGAACCTCTATG 659
Db 60 GGGGAGAGACGTACTCTGAGGCGAGGAGTGCCAGGTAAGCGCGAGCCTTCTTACG 119
Qy 660 CCGCTTCTCCACCGCGCGCTTCCACCTGCGCGCTTGGCGTGTGCGGTTGTCTTACC 719
Db 120 CCGTGTCTCCACCGTAGCGCTTCTACCTGCGCTCTGTGTGTGTCTCTCGGTACT 179
Qy 720 GGAAGATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGTGCGCGT 779
Db 180 GGAAGATCTAAGGCTCCAGGTTCCGCTGGGTCTCCAGGAACCAATACGCT- ---- 234
Qy 780 TGCCGCGCACCATTCAGGTGAGGTCCAAGGTAAAGAGACCTTGATGAGGTGAAGTGG 839
Db 235 ----CTCACCATATCCGAAGCTGTGGAGGTCAAGACTCTGCCAAACAGCCAGATGG 290
Qy 840 TGTTCACGGCAATTGCAAGCAACGTTGCTCTCCAGGTGAGCGGGAAGTCTCTGGCGGG 899
Db 291 TGTTCACGGTCC---GCCACGCCACCGTCACCTTCAGCCAGAGAGGACACGTGCGGG 347
Qy 900 AGCAGAGAGGCGGAGCGAGCAGCCATGATGGTGGGAATTCCTGATTCGCGTGTGTGCTGT 959
Db 348 AGCAGAGAGGAGCGGCGCGCCCTCATGGTGGGCACTCTCATTTGGCGGTTCGCTCT 407

QY 960 GCTGATCCCTCTCTTCTGACGGAACATCATACGCCCACTCTGTGCTGCAGCTGCCCC 1019
 Db 408 GCTGATCCCTCTCTTCTCAACGAGCTCATAGTCCCTCTGCTCTGTGACATCCCG 467
 QY 1020 CCACTCTGGAAGCATATTTCTGTGGTGGTCTGCTCAATTTCTTCTCAACCCCTGA 1079
 Db 468 CCACTCTGGAAGCATATTTCTGTGGTGGTCTGCTCAATTTCTTCTTAAACCCCTGA 527
 QY 1080 TTTTACACAGCTTTTAAAC-AAGAACTACAACAAATGCCCTTCAAGAGCTCTTTTACTAAGCAG 1138
 Db 528 TCTATACGCTTTCACGAGAACTACACAGCGCTTCAAGAACTCTTTTCTAGGCAA 587
 QY 1139 AGATGA 1144
 Db 588 CACTGA 593

RESULT 7
 Bx280849
 LOCUS
 DEFINITION Bx280849 539 bp mRNA linear EST 04-MAR-2003
 IMAGE:5173354, mRNA sequence.

ACCESSION Bx280849
 VERSION Bx280849
 KEYWORDS Bx280849.1 GI:28614119
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 539)
 Ebert,L., Heil,O., Hennig,S., Neubert,P., Partsch,E., Peters,M.,
 Radelof,U., Schneider,D. and Korn,B.
 Human UnigeneSet - RZPD3
 Unpublished (2003)
 Contact: Ina Rolfes

RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGP998J111431.
 RZPDLIB; I.M.A.G.B. cDNA Clone Collection;
 Human UnigeneSet - RZPD3 (RZPDLIB No.972)
 http://www.rzpd.de/CloneCards/cgi-
 bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfes
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de

This clone is available royalty-free from RZPD;
 contact RZPD (clone@rzpd.de) for further information. Seq primer:
 M13u, Primer sequence: CATTGTAACGACGCCAGT.

FEATURES
 Location/Qualifiers
 1..539
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGP998J111431 ; IMAGE:5173354"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6, Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dr primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 23.0%; Score 265.4; DB 13; Length 539;
 Best Local Similarity 70.4%; Pred. No. 2.4e-38;
 Matches 388; Conservative 0; Mismatches 151; Indels 12; Gaps 2;

QY 540 TGATCGGCTCGCCCGGTGCGCTCGCGCGCTCATGCGCCTCGCGCGCTGCTCTTTGGCC 599
 Db 1 TGATCGGCTCACTCGGCACTCTCGCTGTCTATCTCTGCGCCCGCTGCTTTTGGCT 60
 QY 600 GGGCGAGGTGTGCGAGCGTCTGGCTCCAGCGTGCAGAGTGAGCGCGGAACCTCTCTATG 659
 Db 61 GGGGAGAGAGTACTCTGAGGGCAGGAGGTGCCAGGTAAAGCGGAGCTTCTTACG 120
 QY 660 CGCCCTTCTCCACCGCGGCGCTTCACTCGCGCTTGGCGTGGCGGTGCTGTGTCTACC 719
 Db 121 CGGTGTCTCCACCGTAGGCGCTTCTACTCGCGCTCTGTGTGTGTCTCTCTGTGTACT 180
 QY 720 GGAAGATCTACGAGGCGGCCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGTCGCGT 779
 Db 181 GGAAGATCTACAGGCTGCAAGTTCGCGTGGCTCCAGGAAGACCAATAGCGT----- 235
 QY 780 TGC CGCGCCACCATGCAAGTGAAGTCAAGTAAAGGAAGCACCTGATGAGGCTGAAGTGG 839
 Db 236 ----CTCACCATATCCGAAGCTGTGAGTGAAGACTCTGCCAAACAGCCCATGAG 291
 QY 840 TGTTCACGCGCATTCGAAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGG 899
 Db 292 TGTTCACGGTCC---GCCACGCCACCGTCACTTCCAGCCAGAAGGGGACACGTGGCGGG 348
 QY 900 AGCAGAAGGAGAGGCGGAGCAGCCATGATGTTGGGAATTCTGATTGCGCGTGTGTGCTGT 959
 Db 349 AGCAGAAGGAGGAGCGGGCGCCCTCATNGTGGGCACTCTCATGCGGTGTGTGCTGTCT 408
 QY 960 GTGGATCCCTCTTCTTCGCGGAACACTCATCAGCCCACTCTGTGCTGCGAGCTGCCCC 1019
 Db 409 GCTGGATCCCTCTTCTTCTCAGCGAGCTCATCAGTCCCTCTGCTCCTGTGACATCCCG 468
 QY 1020 CCATCTGGAAGACATATTTCTGTGGCTTGGCTACTCCAAATCTTCTCAACCCCTGA 1079
 Db 469 CCATCTGGAAGACATATTTCTGTGGCTTGGCTACTCCAAATCTTCTTAAACCCCTGA 528
 QY 1080 TTTTACACAGCT 1090
 Db 529 TCTATACGCT 539
 CENS05MTV 1027 bp DNA linear GSS 01-SEP-2000
 Tetraodon nigroviridis genome survey sequence F3 end of clone
 038018 of library A from Tetraodon nigroviridis, genomic survey
 sequence.
 AL344524.1 GI:8238294
 GSS: genome survey sequence.
 Tetraodon nigroviridis
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontoidea; Tetraodontidae; Tetraodon.
 1
 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizes,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 20296633
 PUBMED 10835645
 2
 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
 Fizes,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)

MEDLINE	20359837
PUBMED	10899143
REFERENCE	3 (bases 1 to 1027)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : secre@genoscope.cns.fr)
COMMENT	- Web : www.genoscope.cns.fr This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon.
FEATURES	Location/Qualifiers
source	1..1027
	/organism="Tetraodon nigroviridis"
	/mol_type="genomic DNA"
	/db_xref="taxon:99883"
	/clone="038018"
	/clone_lib="A"
	/note="Genoscope sequence ID : COAA038BH09AI-end : T3"
ORIGIN	
Query Match	22.2%; Score 255.6; DB 29; Length 1027;
Best Local Similarity	69.2%; Pred.No.1.8e-36;
Matches 426; Conservative	0; Mismatches 171; Indels 19; Gaps 5;
QY	149 CCCTTTCTGTCTTCACGGTGCTCGGTGATGAGCGTGTAGTGTGCTGATCGCTGCCACT 208
DB	351 CCCTTTCCATCTTCAGGTGCTCGCCTCACCTGTGGCCATGCTGGTGGCCACC 410
QY	209 TTCCTGTGGAACCTGTGGTTCGGTCAACATCCCGGGGTCGGTCCCTTCCACCGCGTG 268
DB	411 TTCTGTGTGNACCTGTGTGCTGTGACCATACTGCGTGTGAGGACGTTCACCGGGTG 470
QY	269 CCGCATAACTTGGTGGCTCGACGGCGCTCTCGAACGAACTAGTGGGAGCGCTGGCGATG 328
DB	471 CCCCACAACTGGTAGTTCATGGCCATCTCTCAGCTCATGTGGTGGCGCCCTGGTCATG 530
QY	329 CCACCGAGCCTGGCGAGTGAGCTGTGCACCGGGCGAGTCCGGCTGTGGGCGGAGCGCTG 388
DB	531 CCGTGTAGCCTGTGTCCAGACT---GAA CGGAGGCTGTGGAAGCTCGGCCGGTCTCTC 587
QY	389 TGCCACGTGTGGAPTCTCTTCGACCGCGGAGCCTTCTGTGCTGCCCGCGGCTCGGG 448
DB	588 TGCCAGGTGTGATCTCTTCGAG-----TGCTCTCTGCACGGCCAGCATCTGG 638
QY	449 AACGTGGCGCCATCGCCTGGGCGCGACGGGGCCATCACACGGCACCTTCGACACAG 508
DB	639 AAGGTGAGCGCCAATTGGCGTGGACCGGTACTGGTCCATCACCGAGCACCTCGATFAC 698
QY	509 CTGCGCACCCGCGACCGCGCTCGTTGTCTATGATCGCGCTCGGCCGGGTCCGTCGGCG 568
DB	699 CTGAAGACA CGCAAGAAGATCTCAAATGTGATATCGCCCTCAGTGGTGTGTCTCTCC 758
QY	569 CTCATCGCCTTCGCGCGCTGCTCTTTGGCGGGCGAGGTGTGCGACGCTCGGCTCCAG 628
DB	759 ATCATCTCCTT---GTCCCCCTCTCTCGGCTGGGCGGCGCATCTCGAGGGGATGAAG 815
QY	629 CGCTGCCAGBTGACCGGGAAACCTCTCATGCGCGCTTCTCCACCGCGGGCGCCTTCAC 688
DB	816 ---TSCCAGATGAGCCAGGAGCCGTGCTGATACACCATCTTCTCCACCTTCGGGGCA 872
QY	689 CTGCGGCTTGGCGGTGCGGTTTTGTCTACCGAAGATCTACGAGGGGGCCAAAGTTTCGT 748
DB	873 CTGCGGC-TGTGNTGGTGTCTTTGTCTACTTGAAGATCTACAAAGGCTGCCCAGTTTCG 931
QY	749 TTCGGCGCGCCGCGGA 764
DB	932 ATCGGCTCCGAAAGA 947

RESULT 9
BI603546/c

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BF603546 893 bp mRNA linear EST 07-SEP-2000
603244448F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5287138 5',
mRNA sequence.
BF603546
BF603546.1 GI:15496485
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 893)
NIH-MGC <http://mgs.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb-1@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LHAM11725 row: 0 column: 11
High quality sequence start: 2
High quality sequence stop: 720.

FEATURES
source

1..893
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5287138"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gtcgag); Oligo-dT primed using primer
5'-TTTTTTTTTTTTTTTWN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN

Query Match 22.2%; Score 255.2; DB 12; Length 893;
Best Local Similarity 70.6%; Pred. No. 2e-36;
Matches 373; Conservative 0; Mismatches 143; Indels 12; Gaps 2;

Qy 162 TCACGGTCTCGGTGATCACGCTGCTAGTGTGTCGTCGTCGCACCTTCTCTGGAAC 221
Db 523 TCCTGATACTGTGTTTACTCTTGCTGGCTTTCTGGTGGCGGACGTTTGCCTGGAAC 464

Qy 222 TGCTGGTTCCGGTCACCATCCGGGGTCCGTGCTTCCACCGCGTCCCGCATAACTGG 281
Db 463 TGCTGGTGTGGGACCATCTCTGGTGTAGCACCCTTCCACCGGTGCCCAACACTGG 404

Qy 282 TGGCCTCGACGGCCGTCTCGGACGAACACTAGTGGCAGCGCTGGCGATGCCACGAGCTGG 341
Db 403 TCGCATCATATGGCGCTCTCGATGTCCTGGTGGCGCGCTGGTCATGCCGCTGAGCCTGG 344

Qy 342 CGAGTAGCTGTGACACGGGCGGAGCTCGGCTGCTGGCCGGAGCCTGTGCCAGTGTGGA 401
Db 343 TGACACGAGCTGT---CCTGGCGCGCTGGCAGCTAGGTGCGAGGCTGTGCCAGCTTTGA 287

Qy 402 TCTCTCTGACGCGGAGCGCTGTCTGTGTCGCCCGCGCTCCGGAACGTGGCGGCCA 461
Db 286 TCGCGTGCAGC-----TGCTTTGTGACAGCCAGCATCTGGAACGTGACGCCA 236

Qy 462 TCGCCTTGGGCGCGACGGGGCCATCACAGGACCTTCAGACACACGCTGCACCGCA 521

RESULT 9
BI603546


```

source
1. .583
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 4BOV"
/notes="vector; pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 20 and day 40
embryos."

ORIGIN
Query Match 21.5%; Score 247.6; DB 12; Length 583;
Best Local Similarity 68.0%; Pred. No. 4.3e-35;
Matches 413; Conservative 0; Mismatches 170; Indels 24; Gaps 4;

Qy 322 GCGATGCCACCGAGCTTGGAGTCTCTTCAGCCGGAGCTGTCTGTCTGCCCGCGG 381
Db 1 GGTATGCCGTGAGCTTGTGACGAGTGT---CGGGCGCGCTGGCAGCTGGGCGG 57
Qy 382 GAGCTGTGCCACGTGTGGATCTCTTCAGCCGGAGCTGTCTGTCTGCCCGCGG 441
Db 58 GCGCTGTGCCACGTGTGGATCTGTG-----CTCTGTGCACGGCCAG 108
Qy 442 CTCCGGAACGTGGCGGCATCGCCCTGGCGCGAGCGGGCCATCACACGGCAGCTGCA 501
Db 109 CATCTGAACTGAGCGCCATCGGCTGACCGTACTGTCTCATCACCGTCACTGGA 168
Qy 502 GCACACGCTGGCCACCGCAGCGCGCTCTGTCTCATGATCGGCTTCGCCCGGTGCC 561
Db 169 ATACACACTCGCGCGCGCAGCGCGCTCTCCAACTCATGATCGGCTCACTGGGCGT 228
Qy 562 GTCCGCGCTCATCGCCCTCGCGCGCTCTCTTTGGCGGGCGAGGTGTGGAGCTCG 621
Db 229 CTCCGCGCTCATCTCTTTGGCGCCACTCTCTTCGCTGGGGGGAGACCTACTCAGAGG 288
Qy 622 GCTCAGCGCTGCCAGGTGAGCGGAACCTCTCTATCGCGCTTCTCACCGCGGGCG 681
Db 289 CAGCGAGGTGCCAGGTGAGCGCGAGCCCTCTTACACGTGTCTCACCGTGGGCGC 348
Qy 682 CTTCCACTCGCGCTTGGCGTGGTGCCTTTGTCTTACCGGAAGATCTACGAGGGCGCAA 741
Db 349 CTTCTACCTCGCGCTTGGCGTGGTGTCTTTGTGTACTGGAAGATCTACAAGGGCGCAA 408
Qy 742 GTTCTGTTTCGGCGCGCGCGAGAGCTGTGTCGCCGTTCGGGCCACCATGAGGTAG 801
Db 409 GCTCCGCTGGCGCGCGCGGAAGACCAACAGCGT-----CTCACCCATATCGAAG 459
Qy 802 GTCCAAGTAAAGGAAGCACCTGATGAGGCTGAAGTGTGTTACGGCAGCATTCGAAAGC 861
Db 460 CGTGAGGGGAAGGGCGCGGTTCACGCGCTGAGATGTTCTCGGGCGC---GCCGCGC 516
Qy 862 AACGGTGTCTTCCAGGTGAGCGGAGCTCTTCGGCGGAGCAGAGAGAGCGCAGCAGC 921
Db 517 CAAGTCACTTCCAGGCGCGAGCGCGGAGCGGTGGCGGAGCAGAGAGAGCGCGCGC 576
Qy 922 CATGATG 928
Db 577 CCTCATG 583

RESULT 12
BG715343/c
LOCUS 602677361F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:4799726 5',
DEFINITION mRNA sequence.
ACCESSION BG715343
VERSION BG715343.1 GI:13994526
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE AUTHORS TITLE JOURNAL COMMENT

1 (bases 1 to 472)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapsb@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1AM10689 row: j column: 15
High quality sequence stop: 470.

FEATURES source

1. 472
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4799726"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI
(gtcag); Oligo-dT primed using primer
5'-TTTTTTTTTTTNN-3', size-selected for average
insert size 2.3 kb and normalized to ROT 5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIMH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

ORIGIN

Query Match 20.8%; Score 239.6; DB 12; Length 472;
Best Local Similarity 71.9%; Pred. No. 1.2e-33;
Matches 345; Conservative 0; Mismatches 124; Indels 11; Gaps 2;
Qy 216 GGAACCTGCTGGTCCGGTCAACCATCCCGGGGTCCTTCCACGCGTGGCGGATA 275
Db 472 GGAACCTGCTGGTCCGGTCAACCATCCCGGTAGCGACCTTCCACGCGTGGCGGATA 413
Qy 276 ACTTGTGGCTCGACGCGCGTCTCGAGCACTAGTGGAGCGCTGGCGATGCCACGA 335
Db 412 ACCTGGTGGCATCCATGGCGCGTCTCGATGTCTGTGGCGCGCTGGTCTATGCCGCTGA 353
Qy 336 GCCTGGCGAGTGAAGTGCACCGGGGACGTCGCGTCTGGCGGAGCGCTGTGCCACG 395
Db 352 GCCTGGTGGCAGAGCTGT---CTTGGCGCGCTGGCAGCTAGTTCGAGGCTGTGCC--- 299
Qy 396 TGTGATCTCTTCGACGCGGAGCTGTCTGTCTGCCCGCGCGCTTCGGGAACGTGG 455
Db 298 -----AGCTTTTGGATCGGTGCGAGCTGTCTGTCTGCACGCGCCAGCATCTGGAACGTGA 244
Qy 456 CGGCCATCGCCCTGGCGCGGACCGGGCCATCACAGGCACCTGCAGCACGCTGCCGA 515
Db 243 CGGCCATAGCCCTTACCGCTACTGTGTCCATCAGCGCCACATGGAATACAGCTCCGCA 184
Qy 516 CCGCGACGCGCGCTCGTTGTCTATGTCGCTCGCCCGGGTGCCTGCGCGCTCATCG 575
Db 183 CCGCGAAGTGTCTTCCAACTGATGTCGCGCTCACCTGGGCACTTCCGCTGTCTCT 124
Qy 576 CCCTCGCGCGCTGCTCTTTTGGCGGGGAGAGTGTGCGACGCTCGGCTCCAGCGCTGCC 635
Db 123 CTCTGGCGCGCTGCTTTTGGTGGGAGAGAGCTACTCTGAGGGCAGCGAGGAGTGCC 64
Qy 636 AGGTGAGCGCGGAACCTTCCTATGCGCGCTTCTTCAACCGCGGGCGCTTCCACCTGCCG 695
Db 63 AGGTAAGCGCGGAGCTTCTTCTACGCGGTGTTCTCCACCGTAGGGCGCTTCTACCTGCC 4

Db 655 CCCCCAACACCTTGTATGCTCCATGCGCCATATCTGACGTGATGCTGAGCTTTGGTGATG 596
QY 329 CCACCGAGCTGCGGAGTGAGCTGTCGACCGGGCGACGTGCGCTGCTGGCGCGAGCCTG 388
Db 595 CCTCTCAGTCTGTGTCACGAGCT---GAACGGCCGTGCGTGAAGCTGGTCCGTTCTC 539
QY 389 TGCACGCTGTGGATCTCTTCGACGCGGAGCCTGTCTGTGCTGCCCGCGCCCTCGGG 448
Db 538 TGTCAAGTCTGGATATCGTTGATGTC-----CTGTGCTGTACGGCTAGCAATTGG 488
QY 449 AACGTGGCGCCATGCGCCCTGGSCCGGACGGGGCCATCACAGGCACCTGCGAGCACG 508
Db 487 AACGTGACCGCAATAGCCCTCGACCGTTACTGGTCCATAACTCGACACCTGGAGTACAC 428
QY 509 CTGCGCACCGGACGCGCCTCGTTGCTCATGATCGCGCTCGCCGGGTGCCGTGCGG 568
Db 427 CTAAGACCGGAGAGATCTCCAATGTGATGATCGGGTTAACCTGGCTGCTTTCATCT 368
QY 569 CTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGAGGTGTGCGACGCTCGGCTCCAG 628
Db 367 GTTATTTGCTCTCGCCTCTG---TTTGGCTGGGTGAGACGTATTTCGGAGGAGAACATG 311
QY 629 CGCTGCCAGTGAGCCGGGAACCTCTCTATGCGCCTTCTCCACCGCGCGCCTTCCAC 688
Db 310 GAGTGCCAGGTGAGCCAGAGCGCTCTTACACCATCTTTTCCACATTCGGGGCCTTCTAT 251
QY 689 CTGCGCCTTGGCGTGTGCGCTTGTCTACCGGAAGATCTACGAGCGGCGCAAGTTTCGT 748
Db 250 CTGCGCTCTCTGTGTGTGTGCTGTTGTCTACTGGAAGATTACAAAGCTGCTAAGTTCCGT 191
QY 749 TTGGG 753
Db 190 ATTGG 186

Search completed: September 7, 2004, 17:13:12
Job time : 2870 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 16:25:27 ; Search time 586 Seconds
(without alignments)

9771.633 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence: 1 cgccatggagcgctagcc.....aagcagagatgaacacagg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3267054 seqs, 2485319735 residues

Total number of hits satisfying chosen parameters: 6534108

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq2:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1152	100.0	1152	10	US-09-976-782-15
2	1068	92.7	1150	10	US-09-823-187-9
3	1064.8	92.4	1150	10	US-09-823-187-11
4	1064	92.4	1458	17	US-10-333-946-34
5	1060	92.0	1119	12	US-10-311-671-25
6	1016	88.2	1155	10	US-09-954-342-29
7	756.2	65.6	3086	15	US-10-017-161-697
8	756.2	65.6	3086	16	US-10-292-798-609
9	732.8	63.6	2061	15	US-10-109-532A-1
10	497.6	43.2	1074	15	US-10-225-567A-445
11	497.6	43.2	1074	15	US-10-345-680-39
12	497.6	43.2	1159	15	US-10-345-680-37
13	497.6	43.2	1159	16	US-10-352-684A-15
14	496	43.1	2700	13	US-09-823-245A-617

15	492.8	42.8	1074	11	US-09-826-509-446	Sequence 446, App
16	378.2	32.8	636	9	US-09-750-373-5	Sequence 5, Appli
17	350.4	30.4	568	15	US-10-029-386-9467	Sequence 9467, Ap
18	297	25.8	297	9	US-09-750-373-4	Sequence 4, Appli
19	275.2	23.9	534	15	US-10-029-386-23167	Sequence 23167, A
20	224	19.4	224	15	US-10-029-386-16640	Sequence 16640, A
21	160.2	13.9	293	15	US-10-027-632-213298	Sequence 213298,
22	157.6	13.7	1210	13	US-10-027-632-213298	Sequence 213298,
23	157.6	13.7	1210	16	US-10-027-632-213298	Sequence 213298,
24	157.2	13.6	585	13	US-10-027-632-213299	Sequence 213299,
25	157.2	13.6	585	16	US-10-027-632-213299	Sequence 213299,
26	142.4	12.4	200	15	US-10-109-532A-4	Sequence 4, Appli
27	106	9.2	2826	15	US-10-225-567A-43	Sequence 43, Appli
28	104.4	9.1	1382	15	US-10-101-510-754	Sequence 754, App
29	104.4	9.1	1382	16	US-10-305-720-1256	Sequence 1256, Ap
30	102.8	8.9	1374	15	US-10-001-073-42	Sequence 42, Appl
31	102.8	8.9	1386	15	US-10-001-073-40	Sequence 40, Appl
32	95.2	8.3	1323	15	US-10-345-680-42	Sequence 42, Appl
33	95.2	8.3	1984	15	US-10-225-567A-19	Sequence 19, Appl
34	95.2	8.3	1984	15	US-10-101-510-52	Sequence 52, Appl
35	95.2	8.3	1984	15	US-10-345-680-40	Sequence 40, Appl
36	95.2	8.3	1984	16	US-10-305-720-1077	Sequence 1077, Ap
37	94	8.2	1332	16	US-10-352-684A-1	Sequence 1, Appli
38	94	8.2	1756	16	US-10-305-720-1197	Sequence 1197, Ap
39	94	8.2	2625	9	US-09-954-531-995	Sequence 995, App
40	94	8.2	2625	17	US-10-717-597-32	Sequence 32, Appl
41	93.6	8.1	1323	11	US-09-826-509-448	Sequence 448, App
42	93.6	8.1	1356	9	US-09-993-844-12	Sequence 12, Appl
43	93	8.1	2428	15	US-10-241-313-15	Sequence 15, Appl
44	93	8.1	2455	15	US-10-060-795B-1	Sequence 1, Appli
45	93	8.1	2482	15	US-10-060-795B-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1

US-09-976-782-15
; Sequence 15, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 15
; LENGTH: 1152
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-976-782-15

Query Match 100.0%; Score 1152; DB 10; Length 1152;
Best Local Similarity 100.0%; Pred. No. 7.2e-314; Indels 0; Gaps 0;
Matches 1152; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

QY 1021 CATCTGGAAGCAATATTTCTGTGCTGGCTACTCCAAATCTTTCTTCAACCCCTGAT 1080
DB |||||
1021 CATCTGGAAGCAATATTTCTGTGCTGGCTACTCCAAATCTTTCTTCAACCCCTGAT 1080
DB |||||
1081 TTACACACCTTTTAAACAAGAACTACAATGCCTTCAAGACCTCTTACTAAGCAGAG 1140
DB |||||
1081 TTACACACCTTTTAAACAAGAACTACAATGCCTTCAAGACCTCTTACTAAGCAGAG 1140
DB |||||
1141 ATGAACACAGG 1152
DB |||||
1141 ATGAACACAGG 1152
DB |||||

RESULT 2
US-09-823-187-9
; Sequence 9, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 9
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24) .. (1133)
US-09-823-187-9

Query Match 92.7%; Score 1068; DB 10; Length 1150;
Best Local Similarity 97.7%; Pred. No. 2.9e-290;
Matches 1125; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

QY 1 CGCCATGAGGCGCTAGCCTTTCAAGTGGCCACCGCGCGGTGGCCCTGGGACC 60
DB |||||
20 CGCCATGAGGCGCTAGCCTTTCAAGTGGCCACCGCGCGGTGGCCCTGGGACC 79
DB |||||
61 CGAGACACGAGGACCGGACCCCAAGCCGAGAGGATACGCTTCCGACCCCGAG 120
DB |||||

QY 1 CGCCATGAGGCGCTAGCCTTTCAAGTGGCCACCGCGCGGTGGCCCTGGGACC 60
DB |||||
1 CGCCATGAGGCGCTAGCCTTTCAAGTGGCCACCGCGCGGTGGCCCTGGGACC 60
DB |||||
61 CGAGACACGAGGACCGGACCCCAAGCCGAGAGGATACGCTTCCGACCCCGAG 120
DB |||||
61 CGAGACACGAGGACCGGACCCCAAGCCGAGAGGATACGCTTCCGACCCCGAG 120
DB |||||
121 CGGCGCGCTCTGCGCGGCGGAGGCGCGCTTCTGTCTTCAAGTGGCTGCTGAC 180
DB |||||
121 CGGCGCGCTCTGCGCGGCGGAGGCGCGCTTCTGTCTTCAAGTGGCTGCTGAC 180
DB |||||
181 GCTGTAGTGTGTGATCGCTGCGACCTTTCTGTGGAACCTGTGTTGCGGTCAACAT 240
DB |||||
181 GCTGTAGTGTGTGATCGCTGCGACCTTTCTGTGGAACCTGTGTTGCGGTCAACAT 240
DB |||||
241 CCGCGGGTCCGTGCTTCCACCGGCTGCGCATAACTTGTGTGGCTCGACGCGCTC 300
DB |||||
241 CCGCGGGTCCGTGCTTCCACCGGCTGCGCATAACTTGTGTGGCTCGACGCGCTC 300
DB |||||
301 GGACGAATAGTGGCAGCGCTGGCGATGCCAGGACCTGGCGAGTGTGACACCGG 360
DB |||||
301 GGACGAATAGTGGCAGCGCTGGCGATGCCAGGACCTGGCGAGTGTGACACCGG 360
DB |||||
361 GCGAGCTCGGCTGCTGGGCGGAGCTGTGCCACGTGTGATCTCTTTCGACGCGGAGC 420
DB |||||
361 GCGAGCTCGGCTGCTGGGCGGAGCTGTGCCACGTGTGATCTCTTTCGACGCGGAGC 420
DB |||||
421 CTGTCTGTGCTGCCCGCGGCTCGGAACTGTGGGCGCATCGCCCTGGGCGGAGCG 480
DB |||||
421 CTGTCTGTGCTGCCCGCGGCTCGGAACTGTGGGCGCATCGCCCTGGGCGGAGCG 480
DB |||||
481 GGCCATACACGGACCTTGACACAGCTGCGCACCCCGACCGCGCTCGTGTGTCAT 540
DB |||||
481 GGCCATACACGGACCTTGACACAGCTGCGCACCCCGACCGCGCTCGTGTGTCAT 540
DB |||||
541 GATCGCTGCGCGGGTCCGTGCGGCTGTCATGCGCTCGCGCGCTCTCTTTGGCGG 600
DB |||||
541 GATCGCTGCGCGGGTCCGTGCGGCTGTCATGCGCTCGCGCGCTCTCTTTGGCGG 600
DB |||||
601 GGGCAGGTGTGCGAGCTCGGCTCCAGGCTGCGAGGTGAGCGGACCTCTATGC 660
DB |||||
601 GGGCAGGTGTGCGAGCTCGGCTCCAGGCTGCGAGGTGAGCGGACCTCTATGC 660
DB |||||
661 CGCCTTCTCACCCCGCGGCTTCCACCTGCGGCTTGGCGTGGTGTCTTACCG 720
DB |||||
661 CGCCTTCTCACCCCGCGGCTTCCACCTGCGGCTTGGCGTGGTGTCTTACCG 720
DB |||||
721 GAAGATCTACGAGGCGCAAGTTTGTGTTTGGCGCGCGCGCGAGAGCTGTGCGCGTT 780
DB |||||
721 GAAGATCTACGAGGCGCAAGTTTGTGTTTGGCGCGCGCGCGAGAGCTGTGCGCGTT 780
DB |||||
781 GCGGCCACCATGAGGCTGAGTCCAGGTGAAGGAGCACCTGATGAGGCTGAAGTGT 840
DB |||||
781 GCGGCCACCATGAGGCTGAGTCCAGGTGAAGGAGCACCTGATGAGGCTGAAGTGT 840
DB |||||
841 GTTTCAGGACATGTCAGAGCAACGCTTCTTCCAGGTGAGCGGGAACCTCTGCGGGA 900
DB |||||
841 GTTTCAGGACATGTCAGAGCAACGCTTCTTCCAGGTGAGCGGGAACCTCTGCGGGA 900
DB |||||
901 GAGAGGAGAGCGAGCAGCATGATGTGGGAATCTGATTTGGCGTGTGTTGCTGTG 960
DB |||||
901 GAGAGGAGAGCGAGCAGCATGATGTGGGAATCTGATTTGGCGTGTGTTGCTGTG 960
DB |||||
961 CTGGATCCCTTCTTCCAGGAACTCATGACCCACCTCTGTGCTGACGCTGCCCCC 1020
DB |||||
961 CTGGATCCCTTCTTCCAGGAACTCATGACCCACCTCTGTGCTGACGCTGCCCCC 1020
DB |||||

Db 80 CGAGACCAGCAG-----CGGACCCCCAAGCCGAGAGGGATCTCGGTTCGACCCCGAG 133
QY 121 CGGCGCGTCTTCCGCGGCGGAGCGCGCCCTTCTCTCTTCAAGGTCCTGGTGGTAC 180
Db 134 CGGCGCGTCTTCCGCGGCGGAGCGCGCCCTTCTCTCTTCAAGGTCCTGGTGGTAC 193
QY 181 GCTGCTAGTGTCTGTGATCGCTGCCACATTCTCTGTGGAACCTGTGTTCCGGTCAACAT 240
Db 194 GCTGCTAGTGTCTGTGATCGCTGCCACATTCTCTGTGGAACCTGTGTTCCGGTCAACAT 253
QY 241 CCGCGGGTCCGTGCTTCCACCGCGTGGCGGATAACTTGTGGTCTCGACGGCGTCTC 300
Best Local Similarity 97.5%; Pred. No. 2.3e-289;
Matches 1123; Conservative 0; Mismatches 2; Indels 27; Gaps 3;
QY 1 CGCATGAGGCGCGTAGCTTTCAGTGGCCACCGCGGGTTCCTTCCCTTGGCCCTGGGACC 60
Db 20 CGCATGAGGCGCGTAGCTTTCAGTGGCCACCGCGGGTTCCTTCCCTTGGCCCTGGGACC 79
QY 61 CGAGACAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACCTCGTTCGACCCCGAG 120
Db 80 CGAGACAGCAG-----CGGACCCCAAGCCGAGAGGATACCTCGTTCGACCCCGAG 133
QY 121 CGGCGCGTCTTCCGCGGCGGAGCGCGCCCTTCTCTGTCTTCAAGGTCCTGGTGGTAC 180
Db 134 CGGCGCGTCTTCCGCGGCGGAGCGCGCCCTTCTCTGTCTTCAAGGTCCTGGTGGTAC 193
QY 181 GCTGCTAGTGTCTGTGATCGCTGCCACATTCTCTGTGGAACCTGTGTTCCGGTCAACAT 240
Db 194 GCTGCTAGTGTCTGTGATCGCTGCCACATTCTCTGTGGAACCTGTGTTCCGGTCAACAT 253
QY 241 CCGCGGGTCCGTGCTTCCACCGCGTGGCGGATAACTTGTGGTCTCGACGGCGTCTC 300

RESULT 3

US-09-823-187-11
; Sequence 11, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taulier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 1150
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (24)..(1133)
US-09-823-187-11

Query Match 92.4%; Score 1064.8; DB 10; Length 1150;

Best Local Similarity 97.5%; Pred. No. 2.3e-289;

Matches 1123; Conservative 0; Mismatches 2; Indels 27; Gaps 3;

```

Db 254 CCGCGGCTCGTGCCTTCCACCGCTGCCGATACCTTGTGGCTCGACGGCCGTCTC 313
QY 301 GGACGAATAGTGGCAGCGCTGGCGATGACCCAGCGCTGGCGAGTGGAGTGTCTGACCGG 360
Db 314 GGACGAATAGTGGCAGCGCTGGCGATGACCCAGCGCTGGCGAGTGGAGTGTCTGACCGG 373
QY 361 GCGAGTGGCTGGCGCGGAGCTGTGCCAGCTGTGCCAGTGTGGATCTCTTCGACCGCGGAGC 420
Db 374 GCGAGTGGCTGGCGCGGAGCTGTGCCAGCTGTGGATCTCTTCGACGCC----- 428
QY 421 CTGTCTGTGCTCCCGCGCGCTCGGGAACGTGGCGGCATCGCCCTGGCGCGGAGCG 480
Db 429 ----CTGTGCTGCCCGCGCGCTCGGGAACGTGGCGGCATCGCCCTGGCGCGGAGCG 484
QY 481 GGCATATCACGGCAGCTGTGCAGACACAGTGTGGACCCCGCAGCGCGCTGTGTCTAT 540
Db 485 GGCATATCACGGCAGCTGTGCAGACACAGTGTGGACCCCGCAGCGCGCTGTGTCTAT 544
QY 541 GATCGCTCGCGCGGCTGGCTGGCGCTCATCGCCCTCGCGCGCTGTCTTTGGCG 600
Db 545 GATCGCTCGCGCGGCTGGCTGGCGCTCATCGCCCTCGCGCGCTGTCTTTGGCG 604
QY 601 GCGGAGTGTGTGAGCTCGCTCGCTCCAGCGCTGCCAGTGTAGCCCGGAACCTCTCTATGC 660
Db 605 GCGGAGTGTGTGAGCTCGCTCGCTCCAGCGCTGCCAGTGTAGCCCGGAACCTCTCTATGC 664
QY 661 CGCCTTCTCACCGCGCGCTTCCACTGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 720
Db 665 CGCCTTCTCACCGCGCGCTTCCACTGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 724
QY 721 GAAGATCTACGAGCGCGCAAGTTTCTGTTTCGCGCGCGCGCGAGAGCTGTGCGCTT 780
Db 725 GAAGATCTACGAGCGCGCAAGTTTCTGTTTCGCGCGCGCGCGAGAGCTGTGCGCTT 784
QY 781 GCGCGCACCATGTGAGGTGAGGTCCAAGTAAAGGAAGCACCTGTAGGCTGAAGTGT 840
Db 785 GCGGCGCACCATGC-----AAGTAAAGGAAGCACCTGTAGGCTGAAGTGT 832
QY 841 GTTCAGGACATATGCAAGCAACGCTGTCTTCAGGTGAGCGGAGCTCTTCGCGGA 900
Db 833 GTTCAGGACATATGCAAGCAACGCTGTCTTCAGGTGAGCGGAGCTCTTCGCGGA 892
QY 901 GCAGAGGAGCGAGCAGCAGCATGATGTGGAAATCTGATGGCGCTGTTGTGCTGTG 960
Db 893 GCAGAGGAGCGAGCAGCAGCATGATGTGGAAATCTGATGGCGCTGTTGTGCTGTG 952
QY 961 CTGGATCCCTTCTTCTGACGGAATCATCGCACTCTGTGCTGAGCTGCCCC 1020
Db 953 CTGGATCCCTTCTTCTGACGGAATCATCGCACTCTGTGCTGAGCTGCCCC 1012
QY 1021 CATCTGAAAGCATATTTCTGTGGCTTGGCTTACTTCAATCTTTTCAACCCCTGAT 1080
Db 1013 CATCTGAAAGCATATTTCTGTGGCTTGGCTTACTTCAATCTTTTCAACCCCTGAT 1072
QY 1081 TTACACAGCTTTTAAAGAACTTAAAGAACTTAAAGAACTTAAAGAACTTAAAGAACT 1140
Db 1073 TTACACAGCTTTTAAAGAACTTAAAGAACTTAAAGAACTTAAAGAACTTAAAGAACT 1132
QY 1141 ATGAACACAGGG 1152
Db 1133 ATGAACACAGGG 1144

```

```

RESULT 4
US-10-333-946-34
; Sequence 34, Application US/10333946
; Publication No. US20040023252A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.
; APPLICANT: BURFORD, Neil; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.

```

```

; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Narinder K.
; APPLICANT: HAPLICK, April J.A.; YAO, Monique G.
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.
; APPLICANT: LEE, Ernestine A.; DING, Li
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0176 USN
; CURRENT APPLICATION NUMBER: US/10/333,946
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/23433
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/221,478
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/223,268
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/227,054
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/231,121
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,243
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/232,691
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/235,146
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 34
; LENGTH: 1458
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7476053CB1
US-10-333-946-34

Query Match 92.4%; Score 1064; DB 17; Length 1458;
Best Local Similarity 97.6%; Pred. No. 4.2e-289;
Matches 1121; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

QY 5 ATGAGGCGCTAGAGCTTTTACGTGGCCACCGCGCGCTTGGCCCTTGGCCCTGGGACCCGAG 64
Db 1 ATGAGGCGCGTAGAGCTTTTACGTGGCCACCGCGCGCTTGGCCCTTGGCCCTGGGACCCGAG 60
QY 65 ACCAGCAGCGAGCCCGGACCCCAAGCCGAGAGGAGTACTCGTTTCGACCCCGAGCGGC 124
Db 61 ACCAGCAG-----CGGACCCCAAGCCGAGAGGAGTACTCGTTTCGACCCCGAGCGGC 114
QY 125 GCGGTCTCTCGCGCGCGAGGCGCGCTTCTCTGTCTTTCACGCTCTGTGTGTGAGCGCTG 184
Db 115 GCGGTCTCTCGCGCGCGAGGCGCGCTTCTCTGTCTTTCACGCTCTGTGTGTGAGCGCTG 174
QY 185 CTAGTGTCTCTGATCGCTGCCACTTTCCTGTGGAACTCTGTGTGGAACTCTGTGTGGTACCATCCG 244
Db 175 CTAGTGTCTCTGATCGCTGCCACTTTCCTGTGGAACTCTGTGTGGTACCATCCG 234
QY 245 CCGGTCCGTGCTTCCACCGCGTGGCGATTAAGTGGTGGCTTCGACGCGCGCTCTCGGAC 304
Db 235 CCGGTCCGTGCTTCCACCGCGTGGCGATTAAGTGGTGGCTTCGACGCGCGCTCTCGGAC 294
QY 305 GAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGTGTGAGTGTGTGAGCCGGA 364
Db 295 GAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGTGTGAGTGTGTGAGCCGGA 354
QY 365 CGTGGCTCTGCGCGCGAGCGCTGTGCCAGCTGTGCCAGCTGTGATCTCTTTCGACCGCGGAGCTGT 424
Db 355 CGTGGCTCTGCGCGCGAGCGCTGTGCCAGCTGTGCCAGCTGTGATCTCTTTCGACCG 405
QY 425 CTGTGCTGCCCGCGCGCTCGGGAACGTGGGCGCATCGCCCTGGCGCGCGAGCGGCGCC 484
Db 406 CTGTGCTGCCCGCGCGCTCGGGAACGTGGGCGCATCGCCCTGGCGCGCGAGCGGCGCC 465

```


Qy 485 ATACAGGACCTGACGACACAGCTGCGACCCCGAGCGCGCTCGTTGCTCATGATC 544
Db 466 ATACAGGACCTGACGACACAGCTGCGACCCCGAGCGCGCTCGTTGCTCATGATC 525
Qy 545 GCGCTCGCCGGGTCGCGCTGCGGCTCATGCGCCCTCGCGCGCTGCTTTGGCGGGGC 604
Db 526 GCGCTCGCCGGGTCGCGCTGCGGCTCATGCGCCCTCGCGCGCTGCTTTGGCGGGGC 585
Qy 605 GAGTGTGCGACGCTCGGCTCAGGCTGCGAGCTGAGCGCGGAAACCTCTATGCGGCC 664
Db 586 GAGTGTGCGACGCTCGGCTCAGGCTGCGAGCTGAGCGCGGAAACCTCTATGCGGCC 645
Qy 665 TTCTCCACCCGCGCGCTTCCACCTGCGGCTTGGCGTGGTGGCTTTGTACCGAAG 724
Db 646 TTCTCCACCCGCGCGCTTCCACCTGCGGCTTGGCGTGGTGGCTTTGTACCGAAG 705
Qy 725 ATCTACGAGCGGCAAGTTTCGTTTCGCGCCCGCGCGAGAGCTGCTGCGGTTGCGG 784
Db 706 ATCTACGAGCGGCAAGTTTCGTTTCGCGCCCGCGCGAGAGCTGCTGCGGTTGCGG 765
Qy 785 GCCACCATGAGGTGAGGTCCAGGTAAAGGAAGCACTGATGAGGCTGAAGTGGTTC 844
Db 766 GCCACCATGAGGTGAGGTCCAGGTAAAGGAAGCACTGATGAGGCTGAAGTGGTTC 813
Qy 845 ACGGCACATTGCAAGCAAGCTGCTCTTCCAGGTGAGCGGGGACTCTGCGCGGAGCAG 904
Db 814 ACGGCACATTGCAAGCAAGCTGCTCTTCCAGGTGAGCGGGGACTCTGCGCGGAGCAG 873
Qy 905 AAGGAGGCGGAGCAGCATGATGTTGGGAATTCGTGTTGCGGTGTTGCTGCTGCTGG 964
Db 874 AAGGAGGCGGAGCAGCATGATGTTGGGAATTCGTGTTGCGGTGTTGCTGCTGCTGG 933
Qy 965 ATCCCTTCTTCTGAGGAACTCATGAGCCACTCTGTGCTGAGGCTGCGCCCATC 1024
Db 934 ATCCCTTCTTCTGAGGAACTCATGAGCCACTCTGTGCTGAGGCTGCGCCCATC 993
Qy 1025 TGGAAAAGCATTTTCTGTGCTTGGCTACTCCAAATTTCTTCTCAACCCCTCATTTAC 1084
Db 994 TGGAAAAGCATTTTCTGTGCTTGGCTACTCCAAATTTCTTCTCAACCCCTCATTTAC 1053
Qy 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCCTCTTTACTAAGCAGATGA 1144
Db 1054 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCCTCTTTACTAAGCAGATGA 1113
Qy 1145 ACAGAGG 1152
Db 1114 ACAGAGG 1121

RESULT 5

US-10-311-671-25
; Sequence 25, Application US/10311671
; Publication No. US20040072996A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: LAL, Preeti G.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HAFALIA, April J. A.
; APPLICANT: NGUYEN, Darniel B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YUE, Henry
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LU, Yan
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GRAUL, Richard
; APPLICANT: YAO, Monique G.
; APPLICANT: YANG, Junming

; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: AU-YOUNG, Janice K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: HERMANDEZ, Roberto
; APPLICANT: WALSH, Roderick T.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: HE, Ann
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0131 USN
; CURRENT APPLICATION NUMBER: US/10/311,671
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/US01/19275
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,483
; PRIOR FILING DATE: 2000-06-16
; PRIOR APPLICATION NUMBER: 60/213,954
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 60/215,209
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 60/216,595
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/218,936
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/219,154
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 60/220,141
; PRIOR FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PERL Program
; SEQ ID NO 25
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7474977CB1
US-10-311-671-25

Query Match 92.0%; Score 1060; DB 12; Length 1119;
Best Local Similarity 97.7%; Pred. No. 5.1e-288;
Matches 1114; Conservative 0; Mismatches 5; Indels 21; Gaps 3;
Qy 5 ATGAGGCGCGCTAGCCTTTCAGTGGCCACCGCGCGCTTGCCTTGGCGTGGGACCGGAG 64
Db 1 ATGAGGCGCGCTAGCCTTTCAGTGGCCACCGCGCGCTTGCCTTGGCGTGGGACCGGAG 60
Qy 65 ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGGATCTCGGTTCCGACCCCGAGCGGC 124
Db 61 ACCAGCAG-----CGGACCCCAAGCCCGAGAGGGATCTCGGTTCCGACCCCGAGCGGC 114
Qy 125 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGGTCTCGTGGTGGTACGCTG 184
Db 115 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGGTCTCGTGGTGGTACGCTG 174
Qy 185 CTAGTGTCTGTATCGCTGCCACTTTCTCTGGAACCTGTTCCGGTCAACATCCCG 244
Db 175 CTAGTGTCTGTATCGCGCCACTTTCTCTGGAACCTGTTCCGGTCAACATCCCG 234
Qy 245 CGGTCGCTGCTTCCACCCGCGTCCGCATAACTTGGTGGCCTCGACCGCGCTCTCGGAC 304
Db 235 CGGTCGCTGCTTCCACCCGCGTCCGCATAACTTGGTGGCCTCGACCGCGCTCTCGGAC 294
Qy 305 GAAGTGTGCGAGCGTGGCGATGCCACCGAGCGCTGGCGAGTGTGTGACGCGGCGGA 364
Db 295 GAAGTGTGCGAGCGTGGCGATGCCACCGAGCGCTGGCGAGTGTGTGACGCGGCGGA 354
Qy 365 CGTGGCTGTGGCGGAGCGCTGTGCCAGTGTGTGATCTCTCTTCGACCGCGGAGCGCTGT 424
Db 355 CGTGGCTGTGGCGGAGCGCTGTGCCAGTGTGTGATCTCTCTTCACG-----TG 405
Qy 425 CTGTGCTGCGCGCGCTCGGGAACGTTGGCGGCCATCGCCCTGGCGCGCGAGCGGCGCC 484


```
Db 419 GTGTGATCTCTTCAGAGCC-----CTGTGCTGCCCGCGGCTCGGAAAGTG 469
Qy 455 GCGGCATCGCTCGGCGGCGGAGCGGCGCATCACGGCACCTCAGCACAGCTGCGC 514
Db 470 GCGGCATCGCTCGGCGGCGGAGCGGCGCATCACGGCACCTCAGCACAGCTGCGC 529
Qy 515 ACCCGACCGCGCTCGTGTGCTCATGATCGCGCTCGCGCGGCTCGCGCTCATC 574
Db 530 ACCCGACCGCGCTCGTGTGCTCATGATCGCGCTCGCGCGGCTCGCGCTCATC 589
Qy 575 GCGCTCGCGCGCTGCTTTTGGCGGCGGAGGCTGCGAGCTCGGCTCGAGCGCTGC 634
Db 590 GCGCTCGCGCGCTGCTTTTGGCGGCGGAGGCTGCGAGCTCGGCTCGAGCGCTGC 649
Qy 635 CAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGGCGCTTCCACCTGCG 694
Db 650 CAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGGCGCTTCCACCTGCG 709
Qy 695 CTTGGCGTGTGCGGCTTTGTCTACCGGAAGATCTACGAGGCGGCGCAAGTTTCGTTTCGCG 754
Db 710 CTTGGCGTGTGCGGCTTTGTCTACCGGAAGATCTACGAGGCGGCGCAAGTTTCGTTTCGCG 769
Qy 755 CCGCGCGGAGAGCTGCTGCTGCGGCTGCGGCGGAGGCTGCGAGGCTGCGAGGTAAG 814
Db 770 CCGCGCGGAGAGCTGCTGCTGCGGCTGCGGCGGAGGCTGCGAGGTAAG-----AAG 817
Qy 815 GAAGCACTGATGAGGCTGAAGTGTGTTACCGGCACATTCGAAGCAACGCTGCTTC 874
Db 818 GAAGCACTGATGAGGCTGAAGTGTGTTACCGGCACATTCGAAGCAACGCTGCTTC 877
Qy 875 CAGGTGAGCGGGAAGCTCTGCGGCGGAGCAGAGGAGGCGGAGCGGAGCGGAGCGG 934
Db 878 CAGGTGAGCGGGAAGCTCTGCGGCGGAGCAGAGGAGGCGGAGCGGAGCGGAGCGG 937
Qy 935 ATTCTGATGCGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 994
Db 938 ATTCTGATGCGGCTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 997
Qy 995 CCATCTGTGCTGCTGAGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAC 1054
Db 998 CCATCTGTGCTGAGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTAC 1057
Qy 1055 TCCAAATCTTTCTTCAAGCGGCTGATTTACAGAGCTTTTAAAGAACTACAAAGTGC 1114
Db 1058 TCCAAATCTTTCTTCAAGCGGCTGATTTACAGAGCTTTTAAAGAACTACAAAGTGC 1117
Qy 1115 TTCAAGAGCTTTTACTAAGCAGAGATGAACAGGG 1152
Db 1118 TTCAAGAGCTTTTACTAAGCAGAGATGAACAGGG 1155
```

RESULT 7

```
US-10-017-161-697
; Sequence 697, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 697
; LENGTH: 3086
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
FEATURE:
; NAME/KEY: source
; LOCATION: (1)..(3086)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(277)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1118)..(2192)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2218)..(2886)
US-10-017-161-697

Query Match 65.6%; Score 756.2; DB 15; Length 3086;
Best Local Similarity 95.4%; Pred.No. 2.4e-202;
Matches 810; Conservative 0; Mismatches 13; Indels 26; Gaps 2;

Qy 1 CGCCATGAGGCGCGCTAGCCTTTTTCAGTGGCCACCGCGGCGTTCGCTTGCCTTGGGACC 60
Db 1798 CGCCATGAGGCGCGCTAGCCTTTTTCAGTGGCCACCGCGGCGTTCGCTTGGGACC 1857
Qy 61 CGAGACCAAGAGC--GAGCCCGGACCCCAAGCCGAGAGGATATCGGTTGACCCCGA 119
Db 1858 CGAGACCAAGAGAGGAGCCCGGACCCCAAGCCGAGAGGATATCGGTTGACCCCGA 1917
Qy 120 GCGGCGCGCTGCTGCGGCGGCGGCGCGCTTCTGCTTTCACGCTGCTGCTGGA 179
Db 1918 GCGGCGCGCTGCTGCGGCGGCGGCGCGCTTCTGCTTTCACGCTGCTGCTGGA 1977
Qy 180 CGCTGTAGTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 239
Db 1978 CGCTGTAGTCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2037
Qy 240 TCCCGCGGCTGCTGCTGCTTCCACCGGCTGCGGAGTAACTTGGTGGCTTCGAGCGGCT 299
Db 2038 TCCCGCGGCTGCTGCTGCTTCCACCGGCTGCGGAGTAACTTGGTGGCTTCGAGCGGCT 2097
Qy 300 CGGAGCACTAGTGGCGGCTGCGGATGCCACGAGCGCTGGGAGTGAAGTGTGAGCG 359
Db 2098 CGGAGCACTAGTGGCGGCTGCGGATGCCACGAGCGCTGGGAGTGAAGTGTGAGCG 2157
Qy 360 GCGGAGCTGCGCTGCTGGCGGCGGAGCTGCGGAGTGTGCTTCTTCCAGCGCGGAG 419
Db 2158 GCGGAGCTGCGCTGCTGGCGGCGGAGCTGCGGAGTGTGCTTCTTCCAGCGCGGAG 2217
Qy 420 CCTGT-----CTGTGCTGCGCGCGGCGCTCGGAAAGTG 454
Db 2218 CCTGTGCGGAGTGTGATCTCTTCCAGCGCTGCTGCTGCGCGCGGCGCTCGGAAAGTG 2277
Qy 455 GCGGCGCATGCGCTGGCGGCGGAGCGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 514
Db 2278 GCGGCGCATGCGCTGGCGGCGGAGCGGCGGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2337
Qy 515 ACCCGACCGCGCTGCTGCTCATGATCGCGCTGCGCGGCGGCTGCGGCTCATC 574
Db 2338 ACCCGACCGCGCTGCTGCTCATGATCGCGCTGCGCGGCGGCTGCGGCTCATC 2397
Qy 575 GCGCTCGCGCGCTGCTTTTGGCGGCGGAGGCTGCGAGCTCGGCTTCAGCGCTGC 634
Db 2398 GCGCTCGCGCGCTGCTTTTGGCGGCGGAGGCTGCGAGCTCGGCTTCAGCGCTGC 2457
Qy 635 CAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGGCGGCTTCCACCTGCG 694
Db 2458 CAGGTGAGCGGGAACCTCTATGCGCGCTTCTCCACCGCGGCGGCTTCCACCTGCG 2517
Qy 695 CTTGGCGTGTGCGGCTTTGTCTACCGGAAGATCTACGAGGCGGCGGAGTTCGTTTCGCG 754
Db 2518 CTTGGCGTGTGCGGCTTTGTCTACCGGAAGATCTACGAGGCGGCGGAGTTCGTTTCGCG 2577
Qy 755 CCGCGCGGAGAGCTGCTGCTGCGGCTTCCCGGCGGAGTGAAGTGAAGTGAAGTGAAG 814
Db 2578 CCGCGCGGAGAGCTGCTGCTGCGGCTTCCCGGCGGAGTGAAGTGAAGTGAAGTGAAG 2637
```


Db 426 CTCATCTTCCCGCGCGAGCGCCCTCTCTGCTTTCACCGTCTTGTGGTACTCTA 485
QY 185 CTAGTGCTGCTGATCGCTGCGCACTTCTCTGTGGAACCTGCTGTTCCGCTACCACTCCG 244
Db 486 CTGGTGTGCTGATCGCTGCGCACTTCTTATGGAATCTGTAGTTCTGTGACTATCCGT 545
QY 245 CGGGTCCGTGCTTCCACCGGTGCGGATCACTTGTGGCTCGACGCGCTCTCGGAC 304
Db 546 CGCGTCCGCGCTTCCACCGGTGCGGATCACTTGTGGCTCGACGCGCTCTCGGAC 605
QY 305 GAACTAGTGGCAGCCTGCGGATGCGACGAGCCTGCGGAGTGTGACGCGGCGA 364
Db 606 GTCCCTGGTGGCGGTCTCTGTGATGCTCTGAGCCTGTGTGAGCGAGTTGTCCGCTGGCGA 665
QY 365 COTCGGCTGTGGCGCGAGCCTGTGCGAGTGTGATCTCTCTGCGAGCGCGAGCCTGT 424
Db 666 GTTGGCAGCTAGGACAGAGTCTGTGCGCAGTGTGGAATCTCTTTCGAGC-----TG 716
QY 425 CTGTGCTCCCGCGCGCTCGGAACTGCGGCGCATCGCCCTGCGGCGGACGCGGCC 484
Db 717 TTGTGCTGACCGCGCAGCATCTGGAACGTGCGGCGCATCGCCCTGGATCGCTACTGGA 776
QY 485 ATCAGCGGACCTGCGACACGCTGCGACCGCGAGCCTGCGGAGTGTGATGATC 544
Db 777 ATCAGCGGACCTGCGACACGCTGCGACCGCGAGCCTGCGGAGTGTGATGATC 836
QY 545 CGGCTCGCGCGCGGTGCGGCTGCGGCTGATCGCCTCGCGCGCGCTCTCTTGGCGGCGG 604
Db 837 GCGATCACCTGGGCACTGTCGCGCTCATGCTCTGCGCGCGCTCTCTTGGCTGGGCG 896
QY 605 GAGGTGTGCGAGCTCGGCTCCAGCGTCCAGGTGAGCGGGAACCTCTATGCGCGCC 664
Db 897 GAAGCTATGATGCTCGGCTGAGGCTTGCCAGGTGAGCGGAGGCGCTCTATGCTGTC 956
QY 565 TTCTGACCGCGCGGCTTCCAGCTGCGGCTGCGGAGTGTGCGGTGCTGCTACCGGAG 724
Db 957 TTCTCCACCTGCGGAGCCTTCTAGCTGCTCTAGCGGTGCTGCTCTTCTACTGGA 1016
QY 725 ATCTACGAGGCGGCAAGTTCGTTTCGCGCGCGCGCGGAGCTGTGCTCGCGTTCG 784
Db 1017 ATATACAAAGCGCGAGTTTCGATTCGCTGCGAGACGCGCGCGGCTGTTACCGCTTC 1076
QY 785 GCCACCATGAGGTGAGTCCAAAGTAAAGAACACCTGATGAGGCTGAAAGTGTGTC 844
Db 1077 GCCACCAAGC-----AGGCAAGGAAGCACTCGGAGTCTGAGATGTTGTC 1124
QY 845 ACGGCACTTGAAGCAACGCTGCTCTTCCAGGTGAGCGGGAAGTCTGCGGAGAG 904
Db 1125 ACAGCCGTCGCGAGCAACAGTACCTTCCAGAACAGCGGAGACTCTGCGGAGAG 1184
QY 905 AAGGAGGCGGAGCAGCATGATGTTGGGAATCTGATTTGGCGGTGTTGCTGTGCTGG 964
Db 1185 AAGGAGGCGGAGCAGCATGATGTTGGGATCTTGTATGCGGHTTGTGCTTGTGTTG 1244
QY 965 ATCCCTTCTTCTGACGGAACATCATAGCCCACTCTGTGCTGAGCTGCCCCCAATC 1024
Db 1245 ATCCCTTCTTCTGACGAGCTCATCAGCCCGCTCTGTGCTGAGCTGCCCCCAATC 1304
QY 1025 TGGAAAGCATATTTCTGTGCTGCTGCTACTCAATCTTCTTCAACCCCTGATTTAC 1084
Db 1305 TGGAAAGCATATTTCTGTGCTGCTGCTATTTCAATCTTCTTCAACCCCTGATTTAC 1364
QY 1085 ACAGCTTTTAAAGAACTCAACAAATGCTTCAAGAGCCTCTTACTAGCAGATGA 1144
Db 1365 ACTGCTTTTAAAGAACTCAACAAATGCTTCAAGAGCCTCTTACTAGCAGATGA 1424

RESULT 10

US-10-225-567A-445

; Sequence 445, Application US/10225567A

; Publication No. US20030113798A1

; GENERAL INFORMATION:

; APPLICANT: LifeSpan Biosciences

; APPLICANT: Brown, Joseph P.
; APPLICANT: Burmer, Glenna C.
; APPLICANT: Roush, Christine L.
; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 445
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-445

Query Match 43.2%; Score 497.6; DB 15; Length 1074;

Best Local Similarity 70.6%; Pred. No. 1.1e-129;

Matches 729; Conservative 0; Mismatches 279; Indels 24; Gaps 4;

QY 113 ACCCGAGCGCGCTCTGCGGCGGAGGCGCCCTCTCTGTGGAACCTGCTGCTGCTG 172
Db 67 AGCCTCGGCAAGACGACCTGCGCCCGAGCTCGCCCTGCTCTCGGTCTCGGAGTGT 126
QY 173 GTGGTGAAGCTGTAGTCTGCTGATCGCTGCCACTTTCCTGTGGAACCTGCTGCT 232
Db 127 ATTCTACCTTGTGGGCTTCTGTGGCGGCGAGTTCGCTGGAACCTGCTGCTGCTG 186
QY 233 GTACCATTCGCGGCTCGGCTTCCACCGCTGCGCATAACTTGGTGGCTCGACG 292
Db 187 GCGACCATCTCTGCTAGCACCTTCCACCGCTGCGGCAACCACTGCTGGCATTCAT 246
QY 293 GCGCTTCGAGCAAACTAGTGGCAGCGTGGCGATGCCAGCGCTGGCGAGTGTGAG 352
Db 247 GCGCTCTGAGTGTCTGGTGGCGCGCTGTGCTATGCGCTGAGCTGGTGCATGAGCT 306
QY 353 TCGACGCGGCGAGCTCGGCTGCTGGCGGAGCTGTGCCAGCTGTGGATCTCTTCGAC 412
Db 307 T---CCGGCGCGCTGGAGTGGAGGTGTGCCAGCTTGGATCGCGTGGAC 363
QY 413 GCGGAGCTGCTGTGCTGCGCGCGCTCGGAACTGCGGCGCATCGCCCTGGC 472
Db 364 G-----TGCTTGTGCACGCGCAGCATCTGGAACGTGACGCGCATAGCTGGAC 414
QY 473 GCGACGCGGCGCATCACAGCGACCTGAGCAGCAGCTGTGCGCACCGGAGCGCGCT 532
Db 415 CGCTACTGTGCTCATACGCGGCAATGGAATACACGCTCGCGCACCGCGAGTGTCTCC 474
QY 533 TTGCTCATGATCGGCTCGCGCGGCTCGGCTCGGCGCTCATCGCCCTCGCGCGCTGCT 592
Db 475 AACGTATGATCGGCTCACTCGGCGCTCTCCGCTGTCTCTCTGCGCCCGCTGCT 534
QY 593 TTTGGCGGCGAGGTGTGCGAGCTCGGCTCCAGGCTGCGAGGTGAGCGCGGAACCC 652
Db 535 TTTGGCTGGGAGAGAGTACTCTGAGGCGAGCGAGAGTGCAGGTAAAGCGGAGCT 594
QY 653 TCTATCGCGCTTCTCCACCGCGCGCTTCCACTGCGCTTGGCGGTGGTGGTGGTGGT 712
Db 595 TCTACCGCGTGTCTCCACCGTAGGCGCTTCTACCTGCGCTCTGTGTGTGCTCTTC 654
QY 713 GTCTACCGAGATCTACGAGGCGCAAGTTCGTTTCGCGCGCGCGCGGAGCTGTG 772
Db 655 GTGTACTGGAAGATCTCAAGGCTGCGCAAGTTCGCGTGGGCTCCAGGAAGACCAAT 714
QY 773 CTGCGGTTCGCGCGCACCATGCAAGTGTGAGGTCAAGGTAAAGGAGCACCTGATGAG 832
Db 715 GT-----CTCACCATATCGAGCTGTGGAGTGTGAGGACTCTGCCAACAGCCC 765
QY 833 GAAGTGTGTTCAGCGCAATTCGAAAGCAACGCTGTCTTCCAGGTGAGCGGAGCTCC 892
Db 766 CAGATGTTGTACGGTCC---GCCAGCCACCGTCACTTCCAGCGCAGGAAGGAGCACG 822

GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Venkateswarlu, Karicheti
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: UROLOGICAL DISORDERS USING 1435, 559, 34021, 44099, 25278,
; TITLE OF INVENTION: 641, 260, 55089, 21407, 42032, 46656, 62553, 302, 323,
; TITLE OF INVENTION: 12303, 985, 13237, 13601, 18926, 318, 2058 OR 6351 MOLECULES.
; FILE REFERENCE: MP102-012P1RM OMNI
; CURRENT APPLICATION NUMBER: US/10/345,680
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: US 60/349,511
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/360,500
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/365,041
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/374,063
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: US 60/403,468
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: US 60/414,262
; PRIOR FILING DATE: 2002-09-27
; PRIOR APPLICATION NUMBER: US 60/419,986
; PRIOR FILING DATE: 2002-10-21
; PRIOR APPLICATION NUMBER: US 60/423,809
; PRIOR FILING DATE: 2002-11-05
; PRIOR APPLICATION NUMBER: US 60/429,797
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 66
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 1159
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (64)...(1137)
US-10-345-680-37

Query Match 43.2%; Score 497.6; DB 15; Length 1159;
Best Local Similarity 70.6%; Pred. No. 1.1e-129;
Matches 729; Conservative 0; Mismatches 279; Indels 24; Gaps 4;

Qy 113 ACCCGAGCGCGCTCTCGCGGCGAGCGCGCTTCTGCTTCCAGGTCCTG 172
Db 130 AGCTTCGCAAGAGCAGCTCGGCCAGCTCGCCCTGCTCTCGGTTCGGAGTGCTT 189
Qy 173 GTGTGAGCTGTAGTCTCTGATCGCTGCCACTTTCTCTGTGGAACCTGCTGTTCCG 232
Db 190 ATTCTCACCTTGTCTGGGCTTCTGTGTGCGCGAGCTTTCGCTTGAACCTGCTGTTGCTG 249
Qy 233 GTACCATCCCGGGTCCGTGCTTCCACGCGTGGCGGATTAATTGGTGGCTCGAGC 292
Db 250 GCGACCATCTCCGTGTACGACCTTCCACGCGTGGCCCAACCACTTGGTGGCATCCATG 309
Qy 293 GCCGTCTCGAGCAACTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGGAGTGAGCTG 352
Db 310 GCCGTCTCGAGTGTCTGTGTGGCGCGCTGTGTATGCGCTGAGCCTGTGTCATGAGCTG 369
Qy 353 TCGACCGGGGACGTGCGCTCTGCGCGCGAGCTGTGCCACGTGTGGATCTCTTCCGAC 412
Db 370 T---CCGGGCGCGCTGCGACGTAGTGGAGGCTGTGCCAGCTTGGATCGCGTGGAC 426
Qy 413 GCCGGAGCTGTCTGTGTGCGCGCGCTCGGAGCTGGGAGCTGGCGGCGCATCGCCCTG36C 472
Db 427 G-----TGCTTGTGTCACGCGCCAGCATCTGGAACGTGAGCGGCCATAGCCCTGGAC 477
Qy 473 CGGACGGGGCCATCACAGGCACTGCAGCACAGCTGGCACCGGCGAGCGCGCTCG 532
Db 478 CGTACTGGTCCATCACGCGCCATGGAATACAGCTTCGCGCACCGGAGTGCTCTCC 537
Qy 533 TTGCTCATGATCGCGCTCGCGCGGTGCGCTGCGCGCTCATCGCCCTCGCGCGCTGCTC 592

538 AACGTTCATGCGCGCTCACCTGGGCACTCTCCGCTGTCATCTCTCGCGCGCGCTGCTT 597
593 TTTGGCGGGGCGAGGTGTGCAACGCTCGGTCCAGCGCTGCCAGGTGAGCGGGAAACCC 652
598 TTTGGCTGGGAGAGACGCTACTCTGAGGCGACGAGGAGTGCAGGTAAAGCCGAGCCT 657
653 TCCTATGCGCGCTTCTCCACCGCGCGGCGCTTCCACCTGCGCTTGGCGTGTGCGGCTT 712
658 TCCTACGCGCGTCTCTCACCGTAGGCGCTTCTACCTGCGCGCTCTGTGTGGTGTCTTTC 717
713 GTCTACCGGAAGATCTACGAGCGCGCAAGTTTCGTTTCGGCGCGCGCGGAGAGCTGTG 772
718 GTGTACTGGAAGATCTACAAGCTGCCAAGTTCCGCTGGGCTCCAGGAAGACCAATAGC 777
773 CTGCGGTTGCGCGCCACCATGAGGTGAGGTTCCAAAGGTAAGGAAAGACCTGATGAGGT 832
778 GT-----CTCACCATATCCGAAGCTGTGGAGGTGAAGGACTCTGCCAAACAGGCC 828
833 GAAGTGTGTTTCACGGCACATTTGCAAGCAACGCTGTCTTCCAGGTGAGCGGGGACTCC 892
829 CAGATGGTGTTCACGGTCC---GCCACGCCACCGCTCACCTTCCAGCAGAGGGGACAG 885
893 TGGCGGAGCAGAGAGGAGGAGCGAGCAGCCATGATGGTGGGAATTCATTTGGCGTGT 952
886 TGGCGGAGCAGAGGAGGAGCGGCGCGCTCATGGTGGGCACTCTCATTTGGCGTGTTC 945
953 GTGCTGTGCTGATCCCTTCTTCTGACGGAATCATCAGCCCACTGTGCGCTGCGAC 1012
946 GTGCTGTGCTGATCCCTTCTTCTCACCGAGCTCATCAGTCCCTCTGCTCTCTGTGAC 1005
1013 CTGCCCCCATCTGGAAGCATATTTCTGTGGCTTGGCTTCCAAATCTTCTTCTTCAAC 1072
1006 ATCCCGCCATCTGGAAGCATCTTCTGTGGCTTGGCTTCCAACTCTTCTTCTTAAAC 1065
1073 CCCCTGATTTACAGAGCTTTTAAAGAACTACAACATGCTTCAAGAGCCTCTTTTACT 1132
1066 CCCCTGATCTATACGGCTTTCACAGAGACTACACAGCGGCTTCAAGAACTCTTTTCT 1125
1133 AAGCAGAGATCA 1144
1126 AGGCAACACTGA 1137

RESULT 13
US-10-352-684A-15
; Sequence 15, Application US/10352684A
; Publication No. US20030215452A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; APPLICANT: Carroll, Joseph M.
; APPLICANT: Healy, Aileen
; APPLICANT: Weich, Nadine S.
; APPLICANT: Kelly, Louise M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR TREATING
; TITLE OF INVENTION: HEMATOLOGICAL DISORDERS USING 131, 148, 199, 12303, 13906,
; TITLE OF INVENTION: 15513, 17822, 302, 5677, 194, 14393, 28059, 7366, 12212,
; TITLE OF INVENTION: 1981, 261, 12416, 270, 1410, 137, 1871, 13051, 1847, 1849,
; TITLE OF INVENTION: 15402, 340, 10217, 837, 1761, 8990 OR 13249 MOLECULES
; FILE REFERENCE: MP102-019P1RMNMIM
; CURRENT APPLICATION NUMBER: US/10/352,684A
; CURRENT FILING DATE: 2003-01-28
; PRIOR APPLICATION NUMBER: US 60/354,333
; PRIOR FILING DATE: 2002-02-04
; PRIOR APPLICATION NUMBER: US 60/360,258
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/364,476
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/375,626
; PRIOR FILING DATE: 2002-04-26
; PRIOR APPLICATION NUMBER: US 60/386,494
; PRIOR FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/390,965

QY	353	TCGACGGGCGAGCTGCGGTGCTGTGGCGCGAGCCCTGTGCAAGTGTGATCTCTTCGAC	411
DB	611	T----CCGGCGCGCGTGGCAGCTAGGTGGAGGCTGTGCGAGCTTTGGATCGCGTGCAC	667
QY	413	GCCGAGCGTGTCTGTGTGCTGCCCGCCGCTCGGGAACGTGGCGGCCATCGCCCTGGC	472
DB	668	G-----TGCCTTGTGACGGCGAGCATCTGAAAGTGAAGGCCATAGCCCTGGAC	718
QY	473	CGCGACGGGCGCATCACACGGCACTGACGACACAGCTGCGCACCGCAGCGCGCGCTTCG	532
DB	719	CGCTACTGGTCCATCACGCGGCACATGGAATACACGCTCCGCAACCGCAAGTGCGTCTCC	778
QY	533	TTGCTCATGATCGGCTCGCCCGGCTGCGGTGCGCGCTCATGCGCCTCGCGCGGTGCTC	592
DB	779	AAGTCTATGATCGCGCTCACCTGGGCACTCTCGGCTGTCTCTCTGCGCCGCTGCTT	838
QY	593	TTTGGCCGGGCGAGTGTGGACGCTCGGCTCCAGCGCTGCCAGTGCAGCGCGGGAACCC	652
DB	839	TTTGGCTGGGAGAGACGTACTCTGAGGGCAGCGAGGAGTGCAGGTAAAGCCGGAGCCT	898
QY	653	TCCTATGGCGCTTCTCCACCGCGGCGCCTTCCACTGCGCTTGGCGTGGTCCGCTTT	712
DB	899	TCCTACGCGCTGTTCTCCACCGTAGGCGCCTCTTACCTGCGCTCTGTGTGCTCTCTC	958
QY	713	GTCCTACCGGAAGACTACGAGCGCGCAAGTTTGTGTCGCGCGCGCCGCGAGAGCTGTG	772
DB	959	GTGTACTGGAAGACTCAAGAGCTGCCAAGTTCGCGTGGGCTCCAGGAAGACCAATAGC	1018
QY	773	CTGCGCTTGC CGGCCACCATGCAGGTGAGGTCCAAAGTAAAGGAAGCACCTGATGAGCT	832
DB	1019	GT-----CTACCCATATCCGAAGCTGTGGAGGTGAGGACTCTGCCAAACAGCCC	1069
QY	833	GAAGTGTGTTCAGCGCATTCGCAAGCAACGGTGTCTTTCAGGTGAGCGGGACTCC	892
DB	1070	CAGATGTGTGTTACGGTCC---GCCAGCCACCGTCACTTCCAGCCAGAAGGGGACAG	1126
QY	893	TGGCGGAGCAGAAAGGAGAGCGCAGCAGCATGATGTGGGAATTCGATTGGCGTGTTC	952
DB	1127	TGGCGGAGCAGAAGAGAGCAGCGGCGCGCCTCAATGTGGGCATCTCATTTGGCGTCTC	1186
QY	953	GTGCTGTCTGGATCCCTTTCTTCTGACGSAATCATACGCCACACTCTGTGCTGTGACG	1012
DB	1187	GTGCTGTCTGGATCCCTTTCTTCTCAGCGAGCTCATAGTCCCTCTGCTCTGTGAC	1246
QY	1013	CTGCCCCCATCTGGAAGAAGATTTCTGTGGCTTGGCTACTCCAAATCTTTCTTCAAC	1072
DB	1247	ATCCCCGCCATCTGGAAGAAGCATCTTCTGTGGCTTGGCTACTCCAACTCCTTTTAAAC	1306
QY	1073	CCCTGATTTACACAGCTTTTAAACAAGACTACAAACAATGCCCTTCAAGAGCCTCTTTACT	1132
DB	1307	CCCTGATCTATACGGCTTTCAACAAGACTACACAGCGGCTTCAAGAACTCTTTTCT	1366
QY	1133	AAGCAGAGATGA	1144
DB	1367	AGGCAACACTGA	1378

```

RESULT 15
US-09-826-509-446
; Sequence 446, Application US/09826509
; Publication No. US20030204073A1
; GENERAL INFORMATION:
; APPLICANT: Lehmann-Bruinsma, Karin
; APPLICANT: Liaw, Chen W.
; APPLICANT: Lin, I-Lin
; TITLE OF INVENTION: No. US20030204073A1-Endogenous, Constitutively Activated Known G
; TITLE OF INVENTION: Protein-Coupled Receptors
; FILE REFERENCE: AREN-207
; CURRENT APPLICATION NUMBER: US/09/826,509
; PRIORITY FILING DATE: 2001-04-05
; PRIORITY FILING DATE: 60/195,747
; PRIORITY FILING DATE: 2000-04-07
; PRIORITY FILING DATE: 09/170,496

```

```

; PRIOR FILING DATE: 1998-10-13
; NUMBER OF SEQ ID NOS: 589
; SOFTWARE: PatentIn Version 2.1
; SEQ ID NO 446
; LENGTH: 1074
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-826-509-446

```

Query Match	42.8%	Score 492.8;	DB 11;	Length 1074;
Best Local Similarity	70.3%	Pred. No. 2.4e-128;		
Matches 726;	Conservative	0;	Mismatches 282;	Indels 24;
				Gaps 4;

QY	113	ACCCGAGGGCGCGCTCTGCGGGCCGAGGGCGCCCTTCTCTGTCTTCAAGTCTCG	173
Db	67	AGCTTCGGCAAGACGACCTGCGCCCGAGCTGCGCCCTGCTCTCGGTCTTCGGAGTCTT	126
QY	173	GTGGTGAAGCTGTAGTGTCTGCTGATCGCTGCCACTTTCTCTGTGGAACTGCTGTGTCG	232
Db	127	ATTCTCACCTTGTGTGGCTTCTGTGGTGGCGGACGTTGCTCTGGAACTGCTGTGTGCTG	186
QY	233	GTCAACATCCGCGGGTCCGTGCTTCCACCGGTGCGCATAACTTGTGGTGGCTCGACG	292
Db	187	GGGACCATCTCGTGTACGCACCTTCCACCGGTGCGCCACAACTGTGTGGCATCCATG	246
QY	293	GCGCTCTCGCAAGAACTAGTGGCAGCGCTGGCGATGCCACGAGCCTGGCGAGTGAAGCTG	352
Db	247	GCGCTCTCGGATGTCCTGTGTGGCGCGCTGGTTCATGCGCGTGAAGCTGTGGCATGAGCTG	306
QY	353	TGACCGGGCGGACGTGTGGCTGTGTGGGCGGAGCCTGTGCCACGTGTGGATCTCTTTCGAC	412
Db	307	T-----CGGGCGCGCTGGCAGTGTGGAGGCTGTGCCAGCTTGTGGATGCGCTGCGAC	363
QY	413	GCGGAGCCTGTCTGTGCTGCCCGCGCTCGGGAAGTGTGGCGGCCATCGCCCTGGGC	472
Db	364	G-----TGTTTGTGACCGCGCAGCATCTGGAAGTGTAGCGGCCATAGCCCTGGAC	414
QY	473	CGGACGGGGCCATCACCGCACCTTCGACGACACACGCTGCGACCCGACCGCGCGCTCG	532
Db	415	CGTACTGTGTCCATCACGCGCACATGGAATACACGCTCCGACCCGCAAGTGGCTCTCC	474
QY	533	TTGCTCATGATCGCGCTCGCGCGGTGCGGTGCGGCTCATCGCCCTCGCGCGGTGCTC	592
Db	475	AAGCTCATGATCGCGCTACCTTGGGACACTCTCGCTGTCACTCTCTGTGCGCCCGCTGCTT	534
QY	593	TTTGGCGCGGGGAGGTGTGCAGCGCTCGCGCTCCAGCGTCCAGGTGACCGGTGACCGGAAACC	652
Db	535	TTTGGCTGGGAGAGAGTACTCTGAGGCGAGGAGGTGCCAGTAAAGCCGCGAGCCT	594
QY	653	TGCTATGCGCCCTTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGGTGTGCGGTTT	712
Db	595	TCTACGCGGTGTCTCCACCGTGTAGCGCTTCTACCTGCGCTCTGTGTGTGTGCTCTTC	654
QY	713	GTCTACCGAAGATCTACGAGGGGCGAAGTTTGTGTTTGGCGCGCGCGCGGAGAGTGTG	772
Db	655	GTGTACTGGAAGATCTACAAGGCTGCAAGTTCCGCGTGGGCTCCAGGAAGACCAATAGC	714
QY	773	CTGCGCGTTCGCGGCCACCATGACAGTGTAGGTCCAAAGTTAAAGGAGACCTGATGAGGCT	832
Db	715	GT-----CTCACCATATCCGAAGCTGTGGAGGTGAAGGACTCTGCCAAACAGGCC	765
QY	833	GAAGTGGTGTTCACGGCACATGCGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACCTCC	892
Db	766	CAGATGGTGTTCACGGTCC---GCCACGCCACCGTCACTTCCAGCCAGAGGGGACACG	822
QY	893	TGCGGGAGCAGAAAGGAGGCGAGCGCCATGATGGTGGGAATCTGATTTGGCGTGTGTT	952
Db	823	TGGCGGGAGCAGAAAGGAGCAGCGGCCAAGCTCATGTGTGGCATCTCATTTGGCGTGTTC	882
QY	953	GTGCTGTGCTGATCCCTTCTTCTGAGGAACTCATACGCCCACTCTGTGCGTGTGACG	1012
Db	883	GTGCTGTGCTGATCCCTTCTTCTCACCGAGTCACTCAGTCCCTCTGCTCTGTGAC	942

QY	1013	CTGCCCCCATCTGGAAAGCATATTCTGTGGCTTGGCTACTCCAAATTCCTTTCTTCAAC	1072
Db	943	ATCCCCGGCATCTGGAAAGCATCTCCGTGGCTTGGCTACTCCAACTCCTTCTTTAAC	1002
QY	1073	CCCCTGATTTACACAGCTTTTAAACAAGAACTACAACAATGCCTTCAAGAGCCTCTTTACT	1132
Db	1003	CCCCTGATCTATACGGCTTTCAACAAGAACTACAACAAGCCTTCAAGAACTTCTTTCT	1062
QY	1133	AAGCAGAGATGA	1144
Db	1063	AGGCAACACTGA	1074

Search completed: September 7, 2004, 18:33:10
Job time : 592 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 7, 2004, 14:14:17 ; Search time 4454 Seconds
(without alignments)
11210.412 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 1152

Sequence: 1 cgccatgagccgtagcc.....aagcagagatgaacacaggg 1152

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sy.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	1152	100.0	1152	6	AX527761	Sequence
2	1068	92.7	1150	6	AX301902	Sequence
3	1064.8	92.4	1150	6	AX301904	Sequence
4	1064	92.4	1458	6	AX375244	Sequence
5	1060	92.0	1119	6	AX342670	Sequence
6	1016	88.2	1155	6	AX526745	Sequence
7	756.2	65.6	1086	6	AX646417	Sequence
8	756.2	65.6	44312	9	AB065453	Homo sapi
9	756.2	65.6	112883	9	AC009404	Homo sapi
10	749.8	65.1	184328	2	AC011638	Homo sapi
11	745	64.7	977	9	HSA308679	Homo sapi
12	732.8	63.6	2036	6	A39678	Sequence 1
13	732.8	63.6	2061	10	MM5HT5BR	Sequence
14	729.8	63.4	2226	6	AR080627	Sequence
15	729.8	63.4	2240	10	RAT5HT5B	Sequence
16	497.6	43.2	1074	6	AX549160	Sequence
17	496	43.1	1074	9	AF498985	Homo sapi
18	494.4	42.9	2911	9	AK094255	Homo sapi
19	492.8	42.8	1074	6	AX280823	Sequence
20	488.2	42.4	207556	10	AC128353	Sequence
21	487.8	42.3	240622	2	AC073793	Mus muscu
22	484.4	42.0	1073	6	A74295	Sequence 7
23	484.4	42.0	1073	6	AR038836	Sequence
24	424.6	36.9	1686	6	A74289	Sequence 1
25	424.6	36.9	1686	6	AR038830	Sequence
26	424.6	36.9	1686	6	AX392847	Sequence
27	424.6	36.9	1686	10	MM5HT5SE	Sequence
28	416	36.1	1954	10	RAT5HT5A	Mus muscu
29	416	36.1	2155	6	AR080628	Sequence
30	394.8	34.3	110000	2	AC073744_0	Mus muscu
31	378.8	32.9	203822	2	AC146825	Otolemur
32	378.2	32.8	636	6	AX188694	Sequence
33	377.2	32.7	194024	2	AC146728	Otolemur
34	350.4	30.4	184328	2	AC011638	Homo sapi
35	348.8	30.3	459	9	HSA308680	Homo sapi
36	334.4	29.0	336	6	A39680	Sequence 3
37	320	27.8	840	9	HS5HT5A1	H. sapiens D
38	320	27.8	840	11	G28709	Sequence
39	318.4	27.6	148832	2	AC074285	Homo sapi
40	318.4	27.6	159041	9	AC093726	Homo sapi
41	316.8	27.5	207644	2	AC146072	Pan trogl
42	316.4	27.5	310072	2	AC111733	Rattus no
43	297	25.8	297	6	AX188693	Sequence
44	273.6	23.8	110000	2	AC111011_2	Continuation (3 of
45	273.6	23.8	205677	2	AC068651	Mus muscu

ALIGNMENTS

RESULT 1
AX527761
LOCUS
DEFINITION Sequence 15 from Patent WO0230974.
ACCESSION AX527761
VERSION AX527761.1 GI:25172284
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
1
Grosse, W.M., Alsobrook, J.P., Lepley, D.M., Burgess, C.E., Mishra, V.,
Kekuda, R., Li, L., Padigaru, M., Shinkets, R.A., Zerkhus, B.D.,
Spytek, K.A., Edinger, S., Gerlach, V., Macdougall, J., Stone, D.,

Pred. No. is the number of results predicted by chance to have a


```
QY 121 CGGCGCGCTCTGCGCGGCGAGGCGCGCCCTTCTCTGCTTTCACGGTCTCTGGTGGTAC 180
DB 134 CGGCGCGCTCTGCGGCGCGAGGCGCGCCCTTCTCTGCTTTCACGGTCTCTGGTGGTAC 193
QY 181 GCTGCTAGTGTCTGATCGCTGCACCTTCTCTGTGGAACCTGCTGGTTCGGTCAACAT 240
DB 194 GCTGCTAGTGTCTGATCGCTGCACCTTCTCTGTGGAACCTGCTGGTTCGGTCAACAT 253
QY 241 CCGCGGGTCCGTGCTCTTCCACCGGCTGCGCCATAACTTGGTGGCTCGACGGCGGCTC 300
DB 254 CCGCGGGTCCGTGCTCTTCCACCGGCTGCGCCATAACTTGGTGGCTCGACGGCGGCTC 313
QY 301 GGACGAACCTAGTGGCAGCGCTGCGGATGCCACCGAGCGCTGCGAGTGTGACCGG 360
DB 314 GGACGAACCTAGTGGCAGCGCTGCGGATGCCACCGAGCGCTGCGAGTGTGACCGG 373
QY 361 GCGACGTGCGTGTGGCGCGAGCGCTGTGCGACGTGTGATCTCTTTCGACGGCGAGC 420
DB 374 GCGACGTGCGTGTGGCGCGAGCGCTGTGCGACGTGTGATCTCTTTCGACGGC----- 428
QY 421 CTGCTGTGCTGCGCGCGCTCGGGAACGTGCGGCCATCGCCCTGGGCGCGAGCG 480
DB 429 -----CTGCTGTGCGCGCGCGCTCGGGAACGTGCGGCCATCGCCCTGGGCGCGAGCG 484
QY 481 GGCCATCACACGCGACCTGACGACACAGCTGCGCACCGCGCGCGCTCTGTGCTCAT 540
DB 485 GGCCATCACACGCGACCTGACGACACAGCTGCGCACCGCGCGCGCTCTGTGCTCAT 544
QY 541 GATCGCGCTCGCGGGTGGTGGCGCTCATCGCCCTGCGCGCGCTGCTCTTTGGCGG 600
DB 545 GATCGCGCTCGCGGGTGGTGGCGCTCATCGCCCTGCGCGCGCTGCTCTTTGGCGG 604
QY 601 GGGCGAGCTGTGCAACCTCGGCTCCAGCGCTGCGAGTGCAGCGCGGAAACCTCTATGC 660
DB 605 GGGCGAGTGTGCAACCTCGGCTCCAGCGCTGCGAGTGCAGCGCGGAAACCTCTATGC 664
QY 661 CGCCTTCTCCACCGCGCGCGCTTCCACCTGCGCGCTTGGGCTGGTGGCGCTTGTACCG 720
DB 665 CGCCTTCTCCACCGCGCGCGCTTCCACCTGCGCGCTTGGGCTGGTGGCGCTTGTACCG 724
QY 721 GAAGATCTACAGCGCGCGAAGTTTGGTTTGGCGCGCGCGCGGAGAGCTGTGCGCGGT 780
DB 725 GAAGATCTACAGCGCGCGAAGTTTGGTTTGGCGCGCGCGCGGAGAGCTGTGCGCGGT 784
QY 781 GCGCGGCACCATGCAAGTGAAGTCCAAAGGTAAAGGAAGCACCTGATCAGGCTGAAGTGT 840
DB 785 GCGCGCCACC-----TCCAAAGGTAAAGGAAGCACCTGATCAGGCTGAAGTGT 832
QY 841 GTTTCAGGCACTTGCAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTTGGCGGGA 900
DB 833 GTTTCAGGCACTTGCAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTTGGCGGGA 892
QY 901 GCGAAGGAGAGGAGCAGCAGCATGATGTTGGGAATTCGATGCGCGTGTGTTGTGCTGTG 960
DB 893 GCGAAGGAGAGGAGCAGCAGCATGATGTTGGGAATTCGATGCGCGTGTGTTGTGCTGTG 952
QY 961 CTGATCCCTCTTCTCTGAGCAACTCATCAGCCCACTGTGCGCTGCGAGCCTGCCCCC 1020
DB 953 CTGATCCCTCTTCTCTGAGCAACTCATCAGCCCACTGTGCGCTGCGAGCCTGCCCCC 1012
QY 1021 CATCTGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTTCACACCCCTGAT 1080
DB 1013 CATCTGAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTTCACACCCCTGAT 1072
QY 1081 TTACAGAGCTTTTAAAGAAGCTTACAAATGCTTCAAGAGCTCTTTTACTAAGAGAG 1140
DB 1073 TTACAGAGCTTTTAAAGAAGCTTACAAATGCTTCAAGAGCTCTTTTACTAAGAGAG 1132
QY 1141 ATGAACACAGGG 1152
DB 1133 ATGAACACAGGG 1144
```

```
RESULT 3
AX301904
LOCUS AX301904 1150 bp DNA linear PAT 30-NOV-2001
DEFINITION Sequence 11 from Patent WO01/4851.
ACCESSION AX301904
VERSION AX301904.1 GI:17382954
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Majumder,K., Spaderna,S.K., Taupier,R.J., Padigaru,M.,
Burgess,C.E., Shinkets,R.A., Spytek,K.A., Liu,X., Patturajan,M. and
Gusev,V.Y.
TITLE Novel proteins and nucleic acids encoding same
JOURNAL Patent: WO 01/4851-A 11 11-OCT-2001;
Curagen Corporation (US)
FEATURES
Location/Qualifiers
1..1150
source /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
24..1136
/note="unnamed protein product"
/codon_start=1
/protein_id="CAD13088.1"
/db_xref="GI:17382955"
/db_xref="REMTREMBL:CAD13088"
/translation="MEASLSVATAGVALAIGALPETSPTSPRGIILGTPSGAVLPGR
GPFPSTLVVLLIATFLNLLVPVTIPRVAFHRVPHNLVASTAVSDELVA
ALAMPDSLASELSTGRRLGRSLCHVWISFDALCCPAGLGNVAIALGRDGAATRL
QHTLFRSRLMIALARVPSALIALAPLLFGRGEVDARLQRCQVRSRPSYAFST
KGAFHLPGLVAPFYRIEAKRFRGRRRAVLPLPATVQKPEADAEVVFVAHCK
ATVSFQVSGDSRQEKERRAAMVMGILIVFLVLPFIFLTELISPLCACSPLPIWKS
IFLWLGYSNFFNPLIYTAFNKNYNNAFKSLFTKQR"
ORIGIN
Query Match 92.4%; Score 1064.8; DB 6; Length 1150;
Best Local Similarity 97.5%; Pred. No. 9.3e-168;
Matches 1123; Conservative 0; Mismatches 2; Indels 27; Gaps 3;
QY 1 CGCCATCGAGCGCGCTAGCTTTTCAGTGGCCACCGCGGGTTCGCCCTTGGCGGACC 60
DB 20 CGCCATCGAGCGCGCTAGCTTTTCAGTGGCCACCGCGGGTTCGCCCTTGGCGGACC 79
QY 61 CGAGACAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACCTCGGTTGACCCCGAG 120
DB 80 CGAGACAGCAG-----CGGACCCCAAGCCGAGAGGATACCTCGGTTGACCCCGAG 133
QY 121 CGGCGCGCTCTGCGCGGCGAGGCGCGCTTCTCTGCTTTCACGGTCTCTGGTGGTAC 180
DB 134 CGGCGCGCTCTGCGCGGCGAGGCGCGCTTCTCTGCTTTCACGGTCTCTGGTGGTAC 193
QY 181 GCTGCTAGTGTCTGATCGCTGCACCTTTCCTGTGGAACCTGCTGTTCCGGTCTCGGTGAC 240
DB 194 GCTGCTAGTGTCTGATCGCTGCACCTTTCCTGTGGAACCTGCTGTTCCGGTCAACAT 253
QY 241 CCGCGGGTCCGTGCTCTTCCACCGGCTGCGCCATAACTTGGTGGCTCGACGGCGGCTC 300
DB 254 CCGCGGGTCCGTGCTCTTCCACCGGCTGCGCCATAACTTGGTGGCTCGACGGCGGCTC 313
QY 301 GGACGAACCTAGTGGCAGCGCTGCGATGCCACCGAGCTGGCGAGTGTGACCGG 360
DB 314 GGACGAACCTAGTGGCAGCGCTGCGATGCCACCGAGCTGGCGAGTGTGACCGG 373
QY 361 GCGACGTGCGTGTGGCGCGAGCGCTGTGCGACGTGTGATCTCTTTCGACGGCGAGC 420
DB 374 GCGACGTGCGTGTGGCGCGAGCGCTGTGCGACGTGTGATCTCTTTCGACGGC----- 428
QY 421 CTGCTGTGCTGCGCGCGCGCTCGGGAACGTGCGGCCATCGCCCTGGGCGCGAGCG 480
DB 429 -----CTGCTGTGCGCGCGCGCTCGGGAACGTGCGGCCATCGCCCTGGGCGCGAGCG 484
```

```
/db_xref="taxon:9606"
/note="Incyte ID No: 7476053CB1"

ORIGIN
Query Match 92.4%; Score 1064; DB 6; Length 1458;
Best Local Similarity 97.6%; Pred. No. 1.2e-167;
Matches 1121; Conservative 0; Mismatches 0; Indels 27; Gaps 3;

QY 5 ATGAGGCGCGTAGCTTTTCAAGTGGCCACCGCCGCGTTCCTTGGCTTGGCCATGGGACCCGAG 64
DB 1 ATGAGGCGCGTAGCTTTTCAAGTGGCCACCGCCGCGTTCCTTGGCTTGGCCATGGGACCCGAG 60
QY 65 ACCAGACGCGACCCGCGACCCAGCCGAGAGGATACCTCGTTCGACCCCGAGCGGC 124
DB 61 ACCAGCAG-----CGGACCCCAAGCCGAGAGGGATACTCGTTCGACCCCGAGCGGC 114
QY 135 GCGGTCTCGCGGCGGAGGCGCGCTTCTGTCTTCAAGTTCCTGCTTCAAGTTCCTGCTGAGCGTG 184
DB 115 GCGGTCTCGCGGCGGAGGCGCGCTTCTGTCTTCAAGTTCCTGCTTCAAGTTCCTGCTGAGCGTG 174
QY 185 CTAGTGTCTGATCGCTGCCACTTTCTGTGGAACTGCTGTTCCGGTCAACCATCCCG 244
DB 175 CTAGTGTCTGATCGCTGCCACTTTCTGTGGAACTGCTGTTCCGGTCAACCATCCCG 234
QY 245 GCGGTCTCGCTTCCACCGCGTCCGATAACTTTGGTGGCCCTCGACGGCCGTCTCGGAC 304
DB 235 GCGGTCTCGCTTCCACCGCGTCCGATAACTTTGGTGGCCCTCGACGGCCGTCTCGGAC 294
QY 305 GAACTAGTGGACCGCTCGGATGCCACCGAGCTGGGAGTGAAGTTCGACCCGGCGA 364
DB 295 GAACTAGTGGACCGCTCGGATGCCACCGAGCTGGGAGTGAAGTTCGACCCGGCGA 354
QY 365 CGTGGCTCTCGCGGCGGAGCTGTGCCACAGTGTGGATCTCTTCCGACCCGAGCGCTGT 424
DB 355 CGTGGCTCTCGCGGCGGAGCTGTGCCACAGTGTGGATCTCTTCCGACCCGAGCGCTGT 405
QY 425 CTGTGCTGCTCGCGGCGCTCGGAAAGTGGGCGCATCGCCCTGGGCGCGGAGCGGCGC 484
DB 406 CTGTGCTGCTCGCGGCGCTCGGAAAGTGGGCGCATCGCCCTGGGCGCGGAGCGGCGC 465
QY 485 ATCACACGCGCTCGACACACGCTGGCAGCCGCGAGCGCGCTGTTGCTCATGATC 544
DB 466 ATCACACGCGCTCGACACACGCTGGCAGCCGCGAGCGCGCTGTTGCTCATGATC 525
QY 545 GCGCTCGCCCGGCGTGGCTCGCGCTCATCGCCCTCGCGCGCTGCTTTCGCGCGGCGC 604
DB 526 GCGCTCGCCCGGCGTGGCTCGCGCTCATCGCCCTCGCGCGCTGCTTTCGCGCGGCGC 585
QY 605 GAGGTGTGCGACGCTCGGCTCGAGCGCTGCGAGGTGAGCCGGGAACTCTTATGCGCGC 664
DB 586 GAGGTGTGCGACGCTCGGCTCGAGCGCTGCGAGGTGAGCCGGGAACTCTTATGCGCGC 645
QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGCTGCTTCTTACCGGAAG 724
DB 646 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGCTGCTTCTTACCGGAAG 705
QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGCTGCGGTTGCGC 784
DB 706 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGCTGCGGTTGCGC 765
QY 785 GGCACCATCGAGTGAAGTCCAAAGTAAAGAGACACCTGATGAGGCTGAAGTGTGTTTC 844
DB 766 GGCACCATCGAGTGAAGTCCAAAGTAAAGAGACACCTGATGAGGCTGAAGTGTGTTTC 813
QY 845 AGCGGACATTGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTCTCTGGCGGAGCAG 904
DB 814 AGCGGACATTGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTCTCTGGCGGAGCAG 873
QY 905 AAGGAGGCGGAGCGGACCATGATGTTGGGAATTCGATTTGGCGTGTGTTGCTGTGCTGG 964
DB 874 AAGGAGGCGGAGCGGACCATGATGTTGGGAATTCGATTTGGCGTGTGTTGCTGTGCTGG 933
QY 965 ATCCCCCTTCTCTGCTGAGGAACTCATACGCCCACTCTGTGCTGAGCGCTGCCCCCATC 1024
```

```
481 GGCATCACAGCGCACCTGACAGCACAGCTGCGCACCCGCGAGCGCGCTGCTGCTCAT 540
DB 485 GGCATCACAGCGCACCTGACAGCACAGCTGCGCACCCGCGAGCGCGCTGCTGCTCAT 544
QY 541 GATCGCGTCCGCGGCGTGGCTGCGGCTCATGCGCCCTCGCGCGCTGCTTGGGCG 600
DB 545 GATCGCGTCCGCGGCGTGGCTGCGGCTCATGCGCCCTCGCGCGCTGCTTGGGCG 604
QY 601 GGGCAGGTGTGCGACGCTCGGCTCGAGGCTGCGAGTGAAGCGGACCTCTATGC 660
DB 605 GGGCAGGTGTGCGACGCTCGGCTCGAGGCTGCGAGTGAAGCGGAACTCTATGC 664
QY 661 GCGCTTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGCTGCTTGTACCG 720
DB 665 GCGCTTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGCTGCTTGTACCG 724
QY 721 GAAATCTACGAGGCGGCAAGTTTCTGTTTCGCGCGCGCGCGAGGCTGCTCGCGTT 780
DB 725 GAAATCTACGAGGCGGCAAGTTTCTGTTTCGCGCGCGCGCGAGGCTGCTCGCGTT 784
QY 781 GCGGCGCACATGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTGAGGCTGAAGTGT 840
DB 785 GCGGCGCACATGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTGAGGCTGAAGTGT 832
QY 841 GTTTCAGCGCACATGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTGCTGCGGGA 900
DB 833 GTTTCAGCGCACATGCAAGCAACGCTGCTTCCAGGTGAGCGGAGCTGCTGCGGGA 892
QY 901 CGAGAAGGAGCGGAGCGGACCATGATGTTGGAAATTCGATTTGGCGTGTGTTGCTGTG 960
DB 893 CGAGAAGGAGCGGAGCGGACCATGATGTTGGAAATTCGATTTGGCGTGTGTTGCTGTG 952
QY 961 CTGGATCCCTTCTTCTGAGCGGACTCATGAGCGGCTGCTGCTGAGCGCTGCGCCC 1020
DB 953 CTGGATCCCTTCTTCTGAGCGGACTCATGAGCGGCTGCTGCTGAGCGCTGCGCCC 1012
QY 1021 CATCTGGAAGAGCATATTTCTGCTGCTTGGCTTACTTCAATTTCTTCAACCCCTCAT 1080
DB 1013 CATCTGGAAGAGCATATTTCTGCTGCTTGGCTTACTTCAATTTCTTCAACCCCTCAT 1072
QY 1081 TTACACAGCTTTTAAACAGAACTTACAAATGCTTCAAGAGCTTCTTAAAGCAGAG 1140
DB 1073 TTACACAGCTTTTAAACAGAACTTACAAATGCTTCAAGAGCTTCTTAAAGCAGAG 1132
QY 1141 ATGAACACAGGG 1152
DB 1133 ATGAACACAGGG 1144
```

```
RESULT 4
AX375244 1458 bp DNA linear PAT 01-MAR-2002
LOCUS
DEFINITION Sequence 34 from Patent WO0210387.
ACCESSION AX375244
VERSION AX375244.1 GI:19169962
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Thornton, M., Patterson, C., Lal, P., Burford, N., Yue, H., Gandhi, A.R.,
Elliot, V.S., Ramkumar, J., Baughn, M.R., Kallick, D.A., Wallia, N.K.,
Hafalla, A.J., Yao, M.G., Lu, Y., Tribouley, C.M., Policky, J.L.,
Kearney, L., Graul, R.C., Warren, B.A. and Ding, L.
G-Protein coupled receptors
Patent: WO 0210387-A 34 07-FEB-2002;
Incyte Genomics, Inc. (US)
FEATURES
source
1..1458
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
```



```
QY 240 TCCCGCGGTCGTCCTTCCACCGCGTCCGCATAACTTGTGGCTTCGACGGCGCTCT 299
Db 2038 TCCCGCGGTCGTCCTTCCACCGCGTCCGCATAACTTGTGGCTTCGACGGCGCTCT 2097
QY 300 CGGACGAATAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGTGACCG 359
Db 2098 CGGACGAATAGTGGCAGCGTGGCGATGCCACCGAGCTGGCGAGTGTGACCG 2157
QY 360 GCGGACGTCGCTGTCGGCGGAGCGCTGTGCCAGTGTGATCTCCCTGACGCGGAG 419
Db 2158 GCGGACGTCGCTGTCGGCGGAGCGCTGTGCCAGTGTGATCTCCCTGACGCGGAG 2217
QY 420 CCTGT-----CTGTGCTGCCCGCGCGCTCGGGAACGTG 454
Db 2218 CCTGTGCCACGTGTGATCTCTTCCACGGCTGTGTGCCCGCGCGCTCGGGAACGTG 2277
QY 455 GCGGCATGCGCTGGCGCGGACGCGGCATCAACGGACCTGACAGCAACGCTGCGC 514
Db 2278 GCGGCATGCGCTGGCGCGGACGCGGCATCAACGGACCTGACAGCAACGCTGCGC 2337
QY 515 ACCCGAGCGCGCTGCTTGTGGCGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGC 634
Db 2398 GCGCTGCGCGCTGCTCTTGTGGCGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGC 2457
QY 635 CAGGTGACCGGGAACCTCTATGCGCGCTTCCACCGCGCGCTTCCACGCTGCC 694
Db 2458 CAGGTGACCGGGAACCTCTATGCGCGCTTCCACCGCGCGCTTCCACGCTGCC 2517
QY 695 CTGCGCTGCTGCGCTTGTCTACCGGAGATCTACGAGGCGGCGCAAGTTTCGTTTGGC 754
Db 2518 CTGCGCTGCTGCGCTTGTCTACCGGAGATCTACGAGGCGGCGCAAGTTTCGTTTGGC 2577
QY 755 CGCGCGCGGAGAGTGTGCTGCTGCTGCGCGTTCGCGCCACCATGAGGTGAGTCCAGGTAAG 814
Db 2578 CGCGCGCGGAGAGTGTGCTGCTGCTGCGCGTTCGCGCCACCATGAGGTGAGGTTGAGG 2637
QY 815 GAAGCACT 823
Db 2638 AACTGCT 2646

RESULT 8
AB065453
LOCUS Homo sapiens gene for seven transmembrane helix receptor,
DEFINITION isolate:CBRC7TM_16.
ACCESSION AB065453
VERSION AB065453.1 GI:21928219
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S.,
Tatsumi,S., Aburatani,H., Asai,K. and Akiyama,Y.
TITLE Genome-wide discovery and analysis of human seven transmembrane
helix receptor genes
JOURNAL Unpublished
AUTHORS Suwa,M.
TITLE Direct Submission
JOURNAL Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research
Center (CBRC), National Institute of Advanced Industrial Science
and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan
(E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/,
Tel:81-3-3599-8080, Fax:81-3-3599-8081)
COMMENT This sequence is a seven transmembrane helix receptor candidate
```

predicted from the whole human genome sequences using our automated system that contains programs of gene finding (GenePredict), sequence search, motif-domain assignment and transmembrane helix prediction. And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (RCAST), University of Tokyo].

FEATURES

source

1. 44312

/organism="Homo sapiens"

/mol_type="genomic DNA"

/isolate="CBRC7TM_16"

/db_xref="taxon:9606"

/chromosome="2"

join(201..1018,43777..44112)

/product="seven transmembrane helix receptor"

/evidence=not_experimental

misc_feature

ORIGIN

Query Match 65.6%; Score 756.2; DB 9; Length 44312;

Best Local Similarity 95.4%; Pred. No. 1.2e-116;

Matches 810; Conservative 0; Mismatches 13; Indels 26; Gaps 2;

QY 1 CGCCATGAGGCGCGCTAGCCTTTCAGTGGCCACCGCGGCGTTCGCCCTGGGACC 60

Db 197 CGCCATGAGGCGCGCTAGCCTTTCAGTGGCCACCGCGGCGTTCGCCCTGGGACC 256

QY 61 CGAGACGAGCAGC-GGACCCCGGACCCCAAGCCCGAGAGGATACTCGGTTTCGACCCCGA 119

Db 257 CGAGACGAGCAGCAGGACCCCGGACCCCAAGCCCGAGAGGATACTCGGTTTCGACCCCGA 316

QY 120 GCGGCGCGTCTCTGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGCTCTCGTGTGA 179

Db 317 GCGGCGCGTCTCTGCGGCGCGAGGCGCGCTTCTCTGTCTTTCACGCTCTCGTGTGA 376

QY 180 CGCTCTAGTGTCTGCTGATGCTGCGACTTTCCTGTGGAACCTGCTGGTTCGGTACCA 239

Db 377 CGCTCTAGTGTCTGCTGATGCTGCGACTTTCCTGTGGAACCTGCTGGTTCGGTACCA 436

QY 240 TCCCGCGGTCGTCCTTCCACCGCGTTCGCATAACTTGTGGCTTCGACGGCGCTCT 299

Db 437 TCCCGCGGTCGTCCTTCCACCGCGTTCGCATAACTTGTGGCTTCGACGGCGCTCT 496

QY 300 CGGACGAATAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGTGACCG 359

Db 497 CGGACGAATAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGTGACCG 556

QY 360 GCGGACGTCGCTGTCGGCGGAGCCTGTGCCACGTCGTGATCTCTTTCGACCGGAG 419

Db 557 GCGGACGTCGCTGTCGGCGGAGCCTGTGCCACGTCGTGATCTCTTTCGACCGGAG 616

QY 420 CCTGT-----CTGTGCTGCCCGCGCGCTTCGGGAACGTG 454

Db 617 CCTGTGCCACGTGTGATCTCTTTCACGCGTGTGCTTCCCGCGCGCTTCGGGAACGTG 676

QY 455 GCGGCATGCGCTGGCGCGGACCGGCGCATCACAGGCACTTCGACGACACGCTCGGC 514

Db 677 GCGGCATGCGCTGGCGCGGACCGGCGCATCACAGGCACTTCGACGACACGCTCGGC 736

QY 515 ACCCGAGCGCGCTGCTGCTCATGATCGCGCTGCCCGGGTGCCTCGGCGCTCATC 574

Db 737 ACCCGAGCGCGCTGCTGCTCATGATCGCGCTGCCCGGGTGCCTCGGCGCTCATC 796

QY 575 GCCTTCGCGCGCTGCTCTTTGGCGGCGAGGTGTGCGACGCTCGGCTTCAGCGGTGC 634

Db 797 GCCTTCGCGCGCTGCTCTTTGGCGGCGAGGTGTGCGACGCTCGGCTTCAGCGGTGC 856

QY 635 CAGGTGAGCGGGAACCTCTTATGCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCC 694

Db 857 CAGGTGAGCGGGAACCTCTTATGCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCC 916

```

695  CTTGCGGTGGTGGCGTTTGTCTACGGAGATCTACGAGCGCGCAAGTTTCGTTTCGGC 754
|||||
917  CTTGCGGTGGTGGCGTTTGTCTACGGAGATCTACGAGCGCGCAAGTTTCGTTTCGGC 976
|||||
755  CGCGCCGGAGAGCTGTCTCGGTGCGCGCCACCATGCGAGTGCACAGTAAAG 814
|||||
977  CGCGCCGGAGAGCTGTCTCGGTGCGCGCCACCATGCGAGTGCAGGCTGGCTGAGG 1036
|||||

QY 815 GAAGCACT 823
DB 1037 AACGTTGCT 1045

RESULT 9
AC009404
LOCUS Homo sapiens BAC clone RP11-28H22 from 2, complete sequence.
AC009404
VERSION AC009404.5 GI:10716633
KEYWORDS HTG
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 112883)
Sulston, J.E. and Waterston, R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
99063792
9847074
REFERENCE 2 (bases 1 to 112883)
AUTHORS Cordes, M., Walker, C., Stoneking, T. and Reitz, L.
TITLE The sequence of Homo sapiens BAC clone RP11-28H22
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 112883)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (21-AUG-1999) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 112883)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (26-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
5 (bases 1 to 112883)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (07-OCT-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Oct 7, 2000 this sequence version replaced gi:9454621.
----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiense@wustl.wustl.edu
----- Summary Statistics
-----
Center project name: H_NH0028H22
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Taten, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://bacpac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-425F6, 200 bp overlap; the clone sequenced to the right is RP11-98C1, 200 bp overlap. Actual start of this clone is at base position 157644 of RP11-425F6; actual end is at base position 21685 of RP11-98C1.

The sequence RP11-28H22 contains imperfect dinucleotide (CT and GC) repeats from base position 53407 to 53507. The region is covered by a single clone; the fidelity of the sequence cannot be guaranteed. The length of sequence is consistent with PCR from clone DNA, and corresponds to restriction digest information: hindIII band size 5664 in silico, and 5694 real.

There are polymorphic base pair differences in the overlaps between the clones RP11-425F6, RP11-28H22, and RP11-425F6.

FEATURES

Location/Qualifiers	Source
1..112883	
/organism="Homo sapiens"	
/mol_type="Genomic DNA"	
/db_xref="taxon:9606"	
/chromosome="2"	
/map="2"	
/clone="RP11-28H22"	
/clone_lib="RPCI-11"	
1..101	misc_feature
/note="similar to EST AW069620 (NID:96024618) cr47g08.xl"	
1..101	misc_feature
/note="similar to EST AW803315 (NID:97855185)"	
8..101	misc_feature
/note="similar to EST AA108891 (NID:91660325) ml56c09.r1"	
56..101	misc_feature
/note="similar to EST AA870919 (NID:92966364) vq28a09.r1"	
644..733	repeat_region
/rpt_family="MER1_type"	
734..807	repeat_region
/rpt_family="MER1_type"	
1062..1194	repeat_region
/rpt_family="Alu"	
2022..2165	misc_feature
/note="similar to EST AW069620 (NID:96024618) cr47g08.xl"	
2022..2165	misc_feature
/note="similar to EST AW803315 (NID:97855185)"	
2034..2165	misc_feature
/note="similar to EST AA108891 (NID:91660325) ml56c09.r1"	
2034..2165	misc_feature
/note="similar to EST AA870919 (NID:92966364) vq28a09.r1"	
2043..2165	misc_feature
/note="similar to EST AW582239 (NID:97257288)"	
2044..2165	misc_feature
/note="similar to EST AA870944 (NID:92966289) vq27b06.r1"	
2044..2165	misc_feature
/note="similar to EST A1021272 (NID:93235608) ub04e12.r1"	
2084..2165	misc_feature
/note="similar to EST AW892944 (NID:98057149)"	
2706..2819	repeat_region
/rpt_family="MER81"	

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 184328)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barna,N., Beckerly,R., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Casile,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., DeRellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M.,
Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J.C., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Holland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Lehoczky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N.,
McEwan,P., McMurk,A., McKernan,K., McLaughlin,J., Meldrim,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (08-OCT-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 1, 2000 this sequence version replaced gi:6957833.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

TITLE
JOURNAL
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
-----Project Information-----
Center project name: L3324
Center clone name: 13_G_16

* NOTE: This is a 'working draft' sequence. It currently
* consists of 55 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.

1 1382: contig of 1382 bp in length
* 1383 1482: gap of 100 bp
* 1483 2509: contig of 1027 bp in length
* 2510 2609: gap of 100 bp
* 2610 3615: contig of 1006 bp in length
* 3616 3715: gap of 100 bp
* 3716 5107: contig of 1392 bp in length
* 5108 5207: gap of 100 bp
* 5208 6331: contig of 1124 bp in length
* 6332 6432: gap of 100 bp
* 6432 7570: contig of 1139 bp in length
* 7571 7670: gap of 100 bp
* 7671 8689: contig of 1019 bp in length
* 8690 8789: gap of 100 bp
* 8790 9889: contig of 1100 bp in length
* 9890 9989: gap of 100 bp
* 9990 11351: contig of 1362 bp in length
* 11352 11451: gap of 100 bp
* 11452 13214: contig of 1763 bp in length
* 13215 13314: gap of 100 bp
* 13315 14472: contig of 1158 bp in length
* 14473 14572: gap of 100 bp
* 14573 16043: contig of 1471 bp in length
* 16044 16143: gap of 100 bp
* 16144 17521: contig of 1378 bp in length
* 17522 17621: gap of 100 bp
* 17622 19059: contig of 1438 bp in length
* 19060 19159: gap of 100 bp
* 19160 20596: contig of 1437 bp in length
* 20597 gap of 100 bp
* 20598 21997: contig of 1301 bp in length
* 21998 22097: gap of 100 bp
* 22098 24537: contig of 2440 bp in length
* 24538 24637: gap of 100 bp
* 24638 25795: contig of 1158 bp in length
* 25796 25896: gap of 100 bp
* 25897 27194: contig of 1299 bp in length
* 27195 27294: gap of 100 bp
* 27295 29425: contig of 2131 bp in length
* 29426 29525: gap of 100 bp
* 29526 31101: contig of 1576 bp in length
* 31102 31201: gap of 100 bp
* 31202 33081: contig of 1880 bp in length
* 33082 33181: gap of 100 bp
* 33182 34843: contig of 1662 bp in length
* 34844 34943: gap of 100 bp
* 34944 37369: contig of 2426 bp in length
* 37370 37469: gap of 100 bp
* 37470 39552: contig of 2083 bp in length
* 39553 39652: gap of 100 bp
* 39653 42320: contig of 2568 bp in length
* 42321 42320: gap of 100 bp
* 42322 44105: contig of 1785 bp in length
* 44106 44205: gap of 100 bp
* 44206 45952: contig of 1747 bp in length
* 45953 46052: gap of 100 bp
* 46053 47615: contig of 1563 bp in length
* 47616 47715: gap of 100 bp
* 47716 51120: contig of 3405 bp in length
* 51121 51220: gap of 100 bp
* 51221 53770: contig of 2550 bp in length
* 53771 53870: gap of 100 bp
* 53871 56963: contig of 3093 bp in length
* 56964 57063: gap of 100 bp
* 57064 59799: contig of 2736 bp in length
* 59800 59899: gap of 100 bp
* 59900 63301: contig of 3402 bp in length
* 63302 63401: gap of 100 bp
* 63402 64177: contig of 776 bp in length
* 64178 64277: gap of 100 bp
* 64278 69030: contig of 4753 bp in length
* 69031 69130: gap of 100 bp
* 69131 73519: contig of 4389 bp in length
* 73520 73619: gap of 100 bp
* 73620 76256: contig of 2637 bp in length
* 76257 76356: gap of 100 bp
* 76357 80598: contig of 4242 bp in length
* 80599 80698: gap of 100 bp
* 80699 84156: contig of 3458 bp in length
* 84157 84256: gap of 100 bp
* 84257 89113: contig of 4857 bp in length
* 89114 89213: gap of 100 bp
* 89214 93633: contig of 4420 bp in length
* 93634 93733: gap of 100 bp
* 93734 98328: contig of 4595 bp in length
* 98329 98428: gap of 100 bp
* 98429 102945: contig of 4517 bp in length
* 102946 103045: gap of 100 bp
* 103046 108095: contig of 5050 bp in length
* 108096 108195: gap of 100 bp
* 108196 114837: contig of 6642 bp in length
* 114838 114937: gap of 100 bp
* 114938 121563: contig of 6626 bp in length
* 121564 121664: gap of 100 bp
* 121664 128405: contig of 6742 bp in length
* 128406 128505: gap of 100 bp
* 128506 133978: contig of 5473 bp in length
* 133979 134078: gap of 100 bp
* 134079 138994: contig of 4916 bp in length
* 138995 139094: gap of 100 bp
* 139095 146606: contig of 7512 bp in length
* 146607 146706: gap of 100 bp


```
Db 666 CGTTGGCAGCTAGGACGAGTCTGTGCCAGCGTGTGATCTCTCTCGACG-----TG 716
QY 425 CTGTGCTCCCGCGGCTCGGAGACCTGCGGCGCATCGCCCTGGCGCGGACGGGCC 484
Db 717 TTGTGCTGACCGCGAGCATCTGGAACCTGCGGCGCATCGCCCTGGATCTGACT 776
QY 485 ATCACAGGACCTGCGACACAGCTGCGCACCCCGCAGCGCGCTCTGTGCTCATGATC 544
Db 777 ATCAGCGGCACTGCACTACAGCTGCGCACCCCGGAGCGGCTCTGTGCTCATGATC 836
QY 545 GCGTGTGCGGCGGCTGCGCTGCGGCGCTCATCGCCCTGCGCGCGCTCTTTTGGCGGGC 604
Db 837 GCGATCACCTGGGCACTGCTGCGGCTCATGCTGCTGCGCCGCTGCTTTTGGCTGGGC 896
QY 605 GAGTGTGCGAGCTGCGCTCCAGGCTGCGGAGTGCAGGAGTGCAGGAGACCTCTCTATGCCGC 664
Db 897 GAAGCCTATGATGTGCGGCTGCGGCTGCGGAGTGCAGGAGTGCAGGAGACCTCTCTATGCTC 956
QY 665 TTCTCCACCGCGCGCTTCCACCTGCGGCTTGGCGTGTGCGGCTTGTCTACCGGAG 724
Db 957 TTCTCCACCTGCGGAGCTTCTACCTGCTCTAGCGTGTGCTCTCTGCTCTACTGGAAA 1016
QY 725 ATCTAAGAGCGGCAAGTTTCTGTTTGGCGCGCGCGGAGAGTGTGCTGCGGTTGCGG 784
Db 1017 ATATACAAAGCGGCAAGTTTCTGATTGCTGCGAGCGGCGGCTGTGATCCGCTTCT 1076
QY 785 GCGACCATGAGGTGAGTCCAGGTAAAGAGACCTGATGAGCTGAAGTGTGTTTC 844
Db 1077 GCCACCGC-----AGCAAAGGAGACCTCGGAGTCTGAGATGTGTTTC 1124
QY 845 ACAGCATTCAAGCAAGCTGTCTTCCAGGTGAGCGGAGTCTCTGCGGAGAGCAG 904
Db 1125 ACAGCGGTGCGGAGCAGTACCTTCCAGACGCGGAGACTCTGCGGAGAGCAG 1184
QY 905 AAGGAGGCGGAGCAGCATTGATGTTGGAAATCTGATTGCGGTGTTTGTGCTGTG 964
Db 1185 AAGGAGGCGGAGCAGCATTGATGTTGGGATCTTGAATGGGTGTTTGTGCTGTG 1244
QY 965 ATCCCTTCTCTGAGGAGTATATAGCGGACTCTGTGCTGAGCTGCGGCGGCGG 1024
Db 1245 ATCCCTTCTCTGAGGAGTATATAGCGGCTCTGTGCTGAGCTGCGGCGGCGG 1304
QY 1025 TGGAAAGCATATTTCTGTGCTGCTGCTTCTTCAATCTTCTTCAACCCCTGATTAC 1084
Db 1305 TGGAAAGCATATTTCTGTGCTGCTGCTTCTTCAATCTGTTCTTCAACCCCTGATTAC 1364
QY 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCTCTTTTACTAAGCAGAGTGA 1144
Db 1365 ACTGCTTTTAAAGAACTCAACAATGCTTCAAGAGCTCTTTTACTAAGCAGAGTGA 1424

RESULT 13
MM5HT5BSR 2061 bp mRNA linear ROD 26-MAY-1993
LOCUS M.musculus mRNA encoding 5-HT5B serotonin receptor.
DEFINITION X69867
ACCESSION X69867.1 GI:288735
VERSION 5-HT5B serotonin receptor.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2061)
Matthes,H., Boschert,U., Amlaiky,N., Grailhe,R., Plassat,J.L.,
Muscatelli,F., Mattei,M.G. and Hen,R.
Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors
define a new family of serotonin receptors: Cloning, functional
expression, and chromosomal localization
Mol. Pharmacol. 43 (3), 313-319 (1993)
JOURNAL 93196607
MEDLINE 8450829
PUBMED
REFERENCE 2 (bases 1 to 2061)
```

```
Matthes,H.
Direct Submission
Submitted (28-DEC-1992) H. Matthes, Laboratoire de Genetique
Moleculaire des Eucaryotes, Departement de Neurobiologie, 11 rue
Humann, F-67000 Strasbourg, FRANCE
Location/Qualifiers
1. .2061
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="brain"
15. .2055
15. .311
312. .1424
/codon_start=1
/product="5-HT5B serotonin receptor"
/protein_id="CAA49501.1"
/db_xref="GI:288736"
/db_xref="GOA:P31387"
/db_xref="SWISS-PROT:P31387"
/translation="MEVSNLSGATPLAAPPGPESCDSPSSGRSMGTPGGLILPGR
EPESAPTVLVLLVLLIATLFLNLLVLVILVRFAHRVPHNLVASTAVSVLVA
VLVPLSLVLSLGRWQLSRCHVWISFDVLCCTASINWVAIALDRWTITRHL
QYTLRTSRASALMIATWALISALIPLLFGWEAYDARLOACQVQSESVAFST
CGAFYLVAVLVFWKIKAAKPRFGRRRRAVPLPATTOAKEAPSESMVFTARR
ATVTFQTSQSWREQEKRAAMVGLILGVFLCNIPFFLTELISPLCACSLLPIWKS
IFLWLGYSNFFNLIYIAFNKNNNNAFKSLFTIKQR"
312. .1088
/label=exon1
1089. .1421
/label=exon2
1425. .2055
exon
exon
5'UTR
ORIGIN
Query Match 63.6%; Score 732.8; DB 10; Length 2061;
Best Local Similarity 79.9%; Pred. No. 1.8e-112;
Matches 911; Conservative 0; Mismatches 202; Indels 27; Gaps 3;
QY 5 ATGAGGCGCTAGCCTTTAGTGGCCACCGCGCGGCTTGCCCTTGGCCTGGGACCGGAG 64
Db 312 ATGGAAGTTTCTAACTCTCAGGCGCCACTCCCGGCTTGCTTCTCCGCGACCTGAG 371
QY 65 ACCAGCAGCGACCGCGGACCCCAAGCCGAGAGGATCTCGGTTCCAGCCCGGACCGC 124
Db 372 AGCTGCAGTAC-----AGCCCAAGTTCCGCGAGGACATGGGATCCACCCAGTGGG 425
QY 125 GCCGTCTCGCGGCGGAGGCGCCCTTCTCTCTTTCACGCTCTGTTGGTGACGCTG 184
Db 426 CTCATCTTGGCGGCGGAGCGCCCTTCTCTGCTTTCACCGTGTGTGTGACTCTA 485
QY 185 CTAGTCTGTGATGCTGCGCACCTTCTGTGGAACCTGCTGGTTCGGGTCAACATCCG 244
Db 486 CTGTTGTGTGATGCTGCTGCACTTCTTATGGAATCTGTAGTCTGTGTAATCTG 545
QY 245 CGGCTCGTGTCTTCCACCGGCTCCGCATAACTTGTGTGGCTCGAGCGGCTCTCGGAC 304
Db 546 CGGCTCGGCTTCCACCGGCTCCGCATAACTTGTGTGGCTCGAGAGCTCTCGGAT 605
QY 305 GAAGTGTGCGGCTGCGGATGCCACCGAGCTGGCGAGTGTGAGCTGTCTGACCGGCGA 364
Db 606 GTCCTGTGCGGCTTCTGTGTGATGCTCTGAGCCCTGCTGAGCGAGTGTCTCGCTGGGCGA 665
QY 365 CGTGGCTGTGGCGGAGCGCTGTGCCAGTGTGATCTCTTTCGAGCGCGGAGCGCTGT 424
Db 666 CGTTGGCAGCTAGGACGAGTCTGTGCCAGCTGTGTGATCTCTCTTCGACG-----TG 716
QY 425 CTGTGCTGCCCGCGGCTCGGAAAGTGTGGCGCATCGCCCTGGCGCGGCGGCGGCC 484
Db 717 TTGTGCTGCAACCGCAGCATCTGGAACGTGGCGCCATCGCCCTGGATCGCTACTGACT 776
QY 485 ATCACAGGACCTGCGACACAGCTCGGACACCGCGAGCGGCTCTGTTGCTCATGATC 544
Db 777 ATCACGGCCACCTGCGAGTACAGCTGCGCACCCCGGAGCGGCTTCTGCGCTCATGATC 836
```


Qy	545	CGCGTCCCGCGGGTGCCTGCGGGCTCATCGCCCTCGCGCGTGTCTTTGGCCCGGGC	604
Db	837	CGGATCACCTGGGCACATGTCGCGCTCATTTGCTCTGCGCCGCTGCTTTTGGCTGGGC	896
Qy	605	GAGGTGTGCAGACGTTCGGCTCCAGCGCTGCCAGTGAAGCCGGGAAACCTTCCTATGCGCC	654
Db	897	GAAGCCATATGATGTTCGGCTGCACGCTTGCCAGTGAAGCCAGAGCCCTCTATGCTGTC	956
Qy	665	TTCTCCACCCGCGCGCTTCCACCTCGCTTGGCGTGTGCGGTGTGCTACCCGAAG	724
Db	957	TTCTCCACCTGCGGAGCTTCTACCTGCCTTAGCGGTGTGCTCTCTCTACTGGA	1016
Qy	725	ATCTACGAGGCGGCAAGTTTCTGTTTCGGCGCGCCGCGAGAGCTGTGCTGCGTTGCCG	784
Db	1017	ATATACAAAGCCGCAAGTTTCGATTTCGTCGACAGCGCGCGGTGTACCGCTTCCT	1076
Qy	785	GCCACCATGCAAGTGAAGTCCAAGTAAAGGAAGCACCTGATGAGCTGAAGTGTGTTTC	844
Db	1077	GCCACCAAGC-----AGGCAAGGAAGACACCTCCGAGCTGAGATGGTGTTC	1124
Qy	845	ACGGCAATTCGAAAGCAAGGTTCTCTTCAGTGAAGCGGGAACCTTCGCGGGAGCAG	904
Db	1125	ACAGCCGTGCGCGAGCAACAGTGACCTTCAGACAGCGAGACCTCTCGCGGAGCAG	1184
Qy	905	AAGGAGAGGCGAGCAGCAAGATGGTGGGAATTCGATGCGCGTGTGTTGCTGTGCTGG	964
Db	1185	AAGGAGAAGCGGCGAGCCATGATGTCGGGATCTTGATGCGCGTGTGTTGCTTGTGG	1244
Qy	965	ATCCGCTTCTTCGACGGAACCTATCAGCCCACTCTGTGCTTCGAGCTGCCCCCATC	1024
Db	1245	ATCCGCTTCTTCGAGCGAGCTCATCAGCCCGCTCTGTGCTGCAGCTGCCCCCATC	1304
Qy	1025	TGGAAGACATATTTCTGTGCTTGGCTACTCCAAATCTTTTTCACACCCCTGATTTAC	1084
Db	1305	TGGAAGACATATTTCTGTGCTTGGATATTCAAATTCGTTCTTCACCCCTGATTTAC	1364
Qy	1085	ACAGCTTTTAAAGAACTACAAAGTCCCTTAAGAGCCCTTTTACTAAGCAGAGATGA	1144
Db	1365	ACTGCTTTTAAAGAAATTAACAATGCTTCAAGAGCCTCTTTTACTAAGCAGAGATA	1424
RESULT 14			
AR080627	2226 bp		
LOCUS	DNA		
DEFINITION	Sequence 1 from patent US 5968817.		
ACCESSION	AR080627		
VERSION	AR080627.1 GI:10007357		
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2226)		
AUTHORS	Surcliffe, J Gregor., Erlander, M.G. and Lovenberg, T.W.		
TITLE	DNA encoding serotonin receptors		
JOURNAL	Patent: US 5968817-A 1 19-OCT-1999;		
FEATURES	Location/Qualifiers		
source	1..2226		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
Query Match	63.4%;	Score 729.8; DB 6; Length 2226;	
Best Local Similarity	79.6%;	Pred. No. 5.5e-112;	
Matches 911; Conservative	0; Mismatches 207; Indels 27; Gaps 3;		
Qy	3	CCATGGAGCGCTAGCTTTTCAGTGGCCACCGCGGTGTGCCCTTGCCTTGGACCCG	62
Db	301	CAATGGAAAGTCTCTAACCTCTCAGCGGCCACCTTCGATTCCTTCCTCGGAGACCG	360
Qy	63	AGACGAGCAGCGGAGCCCGGACCCCAAGCCGAGAGGGATACTCGTTGACCCCGAGC	122
Db	361	AGAGTTCAGTGAC-----ACCCCAAGTTCGGGAGAAGCATGGATCAACCCAGGG	414

RESULT 15										
RAT5HT5B										
LOCUS	2240 bp mRNA linear ROD 26-JUL-1993									
DEFINITION	Rattus norvegicus 5-hydroxytryptamine receptor (5HT5b) mRNA, 5'									
ACCESSION	L10073									
VERSION	L10073.1 GI:310074									
KEYWORDS	5-hydroxytryptamine receptor; serotonin receptor.									
SOURCE	Rattus norvegicus (Norway rat)									
ORGANISM	Rattus norvegicus									
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.									
REFERENCE	1. (bases 1 to 2240)									
AUTHORS	Erlander,M.G., Lovenberg,T.W., Baron,B.M., Lecea,L., Danielson,P.E., Racke,M., Slone,A.L., Siegel,B.W., Foye,P.E., Cannon,K., Burns,J.E. and Sutcliffe,G.J.									
TITLE	Two members of a distinct subfamily of 5-hydroxytryptamine receptors differentially expressed in rat brain									
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 90 (8), 3452-3456 (1993)									
MEDLINE	93234515									
COMMENT	Original source text: Rattus norvegicus (strain Sprague-Dawley) hypothalamus cDNA to mRNA.									
FEATURES	Location/Qualifiers									
source	1..2240									
	/organism="Rattus norvegicus"									
	/mol_type="mRNA"									
	/strain="Sprague-Dawley"									
	/db_xref="taxon:10116"									
	/tissue_type="hypothalamus"									
	303..1415									
	/function="unknown effector coupling"									
	/standard_name="5-hydroxytryptamine receptor"									
	/note="putative"									
	/codon_start=1									
	/label=5-HT5b									
	/product="serotonin receptor"									
	/protein_id="AAA40616.1"									
	/db_xref="GI:310075"									
	/translations="MEVSNLSGATPGIAPPPGSCSDSPSSGSMGSTPGGLILSGR EPFSAFTLVLLVLLIATFLNLLVTLIRVAPRHRVPHNLVASTAVSDVLVA ALNPLSLVSELSAGRWOLGRSLCHWISFDRLCTCTGCCTTGGGACCCG 62 QYTLRRASALMIATWALSALIAPLFGWGRAYDARLQCVQSPSTAVST CAFAYVPLAVLYNKKAFRFRRRRVVLPFPAATKEAPQESSTVFAKRC ATVAFVGDGWSREQEKRAAMVGILGVFLVCWIPPTFLTELVSPLCAGSLPPIMKS IFLWLGYSNSFFNELIYTFAPKNYNNAKSLFTQR"									
ORIGIN										
Query Match	63.4%; Score 729.8; DB 10; Length 2240;									
Best Local Similarity	79.6%; Pred. No. 5.5e-112;									
Matches	911; Conservative 0; Mismatches 207; Indels 27; Gaps 3;									
QY	3	CCATGGAGGCGGTAGCTTTTCAGTGCCACCGCGCGCTTGCCTTGCCTCGGACCCG	62							
Db	301	CAATGGAAGTCTTAACCTCTAGGCGCCACCCCTGGCATGCTTCTTCCTCGGACCCG	360							
QY	63	AGACGACGACGACCGGGACCCGACCCAAAGCCGAGAGGGATACTCGTTGCACCCCGAGCG	122							
Db	361	AGAGCTGCAGTGAC-----AGCCCAAGTTCGCGCAGAAGCATGGGATCACCCAGGGG	414							
QY	123	GCGCCGTCCTCCGGCGGAGCGCGCCCTCTCTGTCCTTCACGGTCTCGTGGTGACGC	182							
Db	415	GGCTCAFTCTGTCGGCGCGGAGCGCGCCCTTCTCTGCTTTCACCGTACTCGTGGTAATC	474							
QY	183	TGCTAGTGTGCTGATCGCTGCCACTTTCCTGTGGAACCTGCTGGTTCGGTCACCATCC	242							
Db	475	TACTGGTGTGCTAGTCGTGCCACTTCTTATGGAATCTGCTAGTCTCGTGACTATCC	534							
QY	243	CGCGGGTCCGTGCTTCCACCGCGTCCGCGATAACTTGGTGGCTCGACGCGCGCTCTCGG	302							
Db	535	TGCGCGTCCGCGCTTCCACCGTGTGCCATAACTTGGTAGCTCGACCGCGCTCTCGG	594							

Search completed: September 7, 2004, 16:25:20
Job time : 4463 secs

QY	303	ACGAACTAGTGGCAGCGCTGGCCGATGCCACCGAGCCTGGCGAGTGACTGTGCACCGGGC	362
Db	595	ACGTCCTGGTGGCGGCTCTGGTGATGCCACTGAGCCTGGTGAGCGAGTGTGCGCTGGCG	654
QY	363	GAGCTCGGCTGCTGGCGCGGAGCCTGTGCCACGTGTGGATCTCTCTTCGACGCGGAGCCT	422
Db	655	GACGTTGGCAGCTGGGCGAGGTCGTGTGCCACGTGTGGATCTCTCTTCGACG-----	705
QY	423	GTCTGTGCTGCCCGCGCGCTCGGAAAGTGTCGCCCATCGCCCTGGGCGCGCAGCGGG	482
Db	706	TGTTGTGTGTCACAGCCAGCATCTGGAACGTGGCGGCATCGCCCTGGATCGCTACTGGGA	765
QY	483	CCATCACAGGCACCTGCAGCACACGCTGGCACCCGCGACCGCGCGCTCGTTGCTCATGA	542
Db	766	CTATCAGCGCCACCTGCAGTACACGCTGGCACCCGCGCGCGCTTCAGCACTTATGA	825
QY	543	TGCGCTTCGCGCGGTGCGGTGCGCGCTCATCGCCCTTCGCGCGCTCTTTTGGCGGG	602
Db	826	TCGCGATCACTGGGCACTGTCCGCCCTCATCGGCTCGCCCGCTGCTTTTGGCTGG	885
QY	603	GCGAGTGTGCGACGCTCGGCTCCAGGCTGCACAGTGAGCGCGGGAACCTCTCATGCGG	662
Db	886	GCGAAGCCTACGATGCTCGGCTCGACGCTGCGAGGTGCGAGTGAGCCAGSAGCCTTCGTACGCCG	945
QY	663	CTTCTCCACCGCGCGGCGCTCCACCTGCGGCTTGCGGTGGTGGCGTTTGTCTACCGGA	722
Db	946	TCTTCTCCACCTGCGAGCTTCTACGTGCTCTGGCGGTGGTCTCTTCGTCTACTGGA	1005
QY	723	AGATCTACGAGCGGCCAAAGTTTCGTTTCGCGCGCGCGGAGAGCTGTGCTGCCGTTGC	782
Db	1006	AGATATACAAAGCGCCAAAGTTTCGATTCGGCGCGAGACGCGCGGCGGTAGTGCCCTGC	1065
QY	783	CGGCGACCATGCAAGTGAAGTCCAGGTAAAGGAAGCACTGATGAGGCTGAAGTGGTGT	842
Db	1066	CGCCACCAACG-----AGGCCAAGGAAGCACCTCAGGAGTCTGAGACGGTAT	1113
QY	843	TCAGGCACATTGCAAGCAAAGTGTCTCTTCAGGTGAGCGGGACTCTCTGCGGGAGC	902
Db	1114	TCACCGCGGTTTCGAGAGCAAGTGGCTTCAGACAAGTGGAGACTCTCTGCGGGAGC	1173
QY	903	AGAAGGAGGCGGAGCAGGCCATGATGGTGGGAATCTGATTGGCGTGTTTGTGCTGTGCT	962
Db	1174	AGAAGGAGAAGCGAGCGCCATGATGGTGGGATCTTGATCGGTGTGTTGTGCTGTGCT	1233
QY	963	GGATCCCTCTTTCCTGACGGAACCTCATCAGCCCACTCTGTGCTGCGAGCCTGCCCCCA	1022
Db	1234	GGATCCCTCTTCTCTGACGAGCTCGTCAGCCCGCTCTCGGCTCGAGCCTGCCACCCA	1293
QY	1023	TCGTGAAAAGCATATTCTGTGGCTGGCTACTCCAAATCTTCTTCAACCCCTCATTT	1082
Db	1294	TCGTGAAAAGCATATTCCTGTGGCTGGGTATTCAAATTCGTTCTTCAATCCCTTAACTCT	1353
QY	1083	ACACGCTTTTAAACAAGAATAACAACAATGCCTTCAAGAGCCTCTTTACTAAGACGAGAT	1142
Db	1354	ACACGSCCTTTAATAGAATAACAACAATGCCTTCAAGAGCCTCTTTACTAAGACGAGAT	1413
QY	1143	GAACA	1147
Db	1414	AAGCA	1418

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:09:52 ; Search time 30 Seconds
(without alignments)

3964.875 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcagagcgccgtagcc.....aagcagagatgaacacaggg 1152

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 399414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 778828

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09976782/runat_31082004_115952_10354/app_query.fasta_1.1351
-DB=Issued Patents AA -QMT=fastan -SUFFIX=n2p.ra1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NOFM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09976782@cgn1_1_46@runat_31082004_115952_10354 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1444.5	67.1	370	2	US-08-031-538-2
2	1022.5	47.5	357	1	US-08-356-405-2
3	1022.5	47.5	357	1	US-08-356-405-9
4	1014.5	47.1	357	2	US-08-031-538-4
5	448	20.8	390	2	US-08-461-812-4
6	448	20.8	390	2	US-08-157-185-15
7	448	20.8	390	3	US-08-281-526B-15
8	448	20.8	390	4	US-09-450-790A-15
9	448	20.8	390	4	US-09-332-837-15
10	448	20.8	390	4	US-09-371-705-4
11	433	20.1	390	1	US-07-817-920-6
12	433	20.1	390	1	US-08-117-006-6

13	433	20.1	390	1	US-08-216-594-6	Sequence 6, Appli
14	433	20.1	390	5	PCT-US93-00149-6	Sequence 6, Appli
15	433	20.1	398	1	US-08-370-542-6	Sequence 6, Appli
16	433	20.1	398	1	US-08-542-358-6	Sequence 6, Appli
17	433	20.1	398	3	US-09-018-351-6	Sequence 6, Appli
18	431.5	20.0	375	1	US-08-370-542-5	Sequence 5, Appli
19	431.5	20.0	375	1	US-08-542-358-5	Sequence 5, Appli
20	431.5	20.0	375	3	US-09-018-351-5	Sequence 5, Appli
21	431.5	20.0	376	1	US-07-817-920-5	Sequence 5, Appli
22	431.5	20.0	376	1	US-08-117-006-5	Sequence 5, Appli
23	431.5	20.0	376	1	US-08-216-594-5	Sequence 5, Appli
24	431.5	20.0	376	5	PCT-US93-00149-5	Sequence 5, Appli
25	430	20.0	377	2	US-08-461-812-2	Sequence 2, Appli
26	430	20.0	377	2	US-08-157-185-14	Sequence 14, Appl
27	430	20.0	377	3	US-08-281-526B-14	Sequence 14, Appl
28	430	20.0	377	4	US-09-450-790A-14	Sequence 14, Appl
29	430	20.0	377	4	US-09-332-837-14	Sequence 14, Appl
30	430	20.0	377	4	US-09-371-705-2	Sequence 2, Appli
31	430	20.0	422	1	US-07-817-920-3	Sequence 3, Appli
32	430	20.0	422	1	US-08-370-542-3	Sequence 3, Appli
33	430	20.0	422	1	US-08-117-006-3	Sequence 3, Appli
34	430	20.0	422	1	US-08-216-594-3	Sequence 3, Appli
35	430	20.0	422	1	US-08-542-358-3	Sequence 3, Appli
36	430	20.0	422	2	US-08-157-185-13	Sequence 13, Appl
37	430	20.0	422	3	US-08-281-526B-13	Sequence 13, Appl
38	430	20.0	422	3	US-09-018-351-3	Sequence 3, Appli
39	430	20.0	422	4	US-09-332-837-13	Sequence 13, Appl
40	430	20.0	422	5	PCT-US93-00149-3	Sequence 3, Appli
41	429	19.9	422	1	US-07-996-772A-12	Sequence 12, Appl
42	428	19.9	422	4	US-09-341-446B-10	Sequence 10, Appl
43	420.5	19.5	421	4	US-09-450-790A-13	Sequence 13, Appl
44	417	19.4	445	2	US-08-157-185-2	Sequence 2, Appli
45	417	19.4	445	3	US-08-281-526B-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-08-031-538-2
; Sequence 2, Application US/08031538
; Patent No. 5968817
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, J Gregor
; APPLICANT: Brlander, Mark G
; APPLICANT: Lovenberg, Timothy W
; TITLE OF INVENTION: HUMAN SEROTONIN RECEPTORS, DNA ENCODING
; TITLE OF INVENTION: THE RECEPTORS, AND USES THEREOF
; NUMBER OF SEQUENCES: 73
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; ADDRESSEE: Patent Counsel
; STREET: 10666 No. 5968817th Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/08/031,538
; APPLICATION NUMBER: 19930315
; FILING DATE: 19930315
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: Fitting, Thomas
; REGISTRATION NUMBER: 34,163
; REFERENCE/DOCKET NUMBER: T5R5099P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-554-2937
; TELEFAX: 619-554-6312
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 370 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-031-538-2

Alignment Scores:

Pred. No.: 1,248-92 Length: 370
 Score: 1444.50 Matches: 295
 Percent Similarity: 82.8% Conservative: 19
 Best Local Similarity: 77.84% Mismatches: 56
 Query Match: 67.06% Indels: 9
 DB: 2 Gaps: 3

US-09-976-782-15 (1-1152) x US-08-031-538-2 (1-370)

```

QY 5 ATGGAGCGCGTAGCTTTTCAGTGGCCACCGCGGCTTCCCTTGGAGCCGAG 64
Db 1 MetGluValCysAsnLeuSerGlyAlaThrProGlyIleAlaPheProProGlu 20
QY 65 ACCAGAGGAGCCGGGACCCAGCCGAGGAGTACTCGTTCGAGCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCGGTCTCTCGCGGCGCGGCGCCCTTCTCTCTTCCACGTCCTCTGCTGAGCGTG 184
Db 39 LeuIleLeuSerGlyArgGluProPheSerAlaPheThrValValValThrLeu 58
QY 185 CTAGTGTCTGATCGCTGCCACTTTCTCTGTGGAACTCTGCTCGGTCCACCTCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78
QY 245 CGGTCTCGTCTTCCACCGGTGCGGATACCTTGGTGGCTCGAGCGCGCTCTCGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACGTAGTGGCGCTGCGGATGCCACCGAGCTGGCGAGTGAGCTGTCGACCGGCGA 364
Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTCTCGCGGCGGAGCTGTGCCACGTGTGGATCTCTCTCGACCGCGAGCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGTCTCGCGCGCGCTCGGAACTGGGCGGCGCATCGCTGGCGGCGCGAGCGGCG 484
Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
QY 485 ATCACACGACCTGACGACACAGCTGCGCACCGCGAGCGCGCTCTGTCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTrpThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCGCGGTCGCGTGGCGCTCATCGCCCTCGCGCGCGCTGTCTTTGCGCGGCG 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGGTGTGACCTCGGCTCCAGCGCTGCCAGGTGAGCGCGGAAACCTCTCATGCGCGC 664
Db 196 GluAlaTyAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTrpAlaVal 215
QY 665 TTCTCCACCGCGCGCTTCCACCGCTGCGTGGCGTGGCTGTGTCTACCGGAG 724
Db 216 PheSerThrCysGlyAlaPheTrpValProLeuAlaValValLeuPheValTrpTrpLys 235
QY 725 ATCTACAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGTGCGGTGCGCG 784
Db 236 IleTyLeuAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
QY 785 GCCACCATGAGGTGAGTCCAGGTAAAGAGACCTCATGAGCTGAGGTGTGTTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271

```

```

QY 845 ACGGCACATTCGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCTGGCGGAGCAG 904
Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGln 291
QY 905 AAGGAGCGGAGCAGCATGATGTGGGAATTCGATTCGGCTGTGTGTGTGTGTGTGTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTCTCTCTGACGGAACCTCATCAGCCACCTCTGTGCTGCTGCTGCTGCTGCTG 1024
Db 312 IleProPhePheLeuThrGluLeuValSerAlaLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAAGCATATTTCTGTGCTTGGCTACTCTCAATTTCTTCAACCCCTGATTTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTrpSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAACAAGTACAACTGCTTCAAGAGCTCTTTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTrpAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

```

RESULT 2

US-08-356-405-2
 ; Sequence 2, Application US/08356405
 ; Patent No. 5807691
 ; GENERAL INFORMATION:
 ; APPLICANT: Amlaiky, No. 5807691rdine
 ; APPLICANT: Boschert, Ursula
 ; APPLICANT: Hen, Rene
 ; APPLICANT: Plassat, Jean-Luc
 ; TITLE OF INVENTION: Polypeptides Having Serotonin Receptor
 ; TITLE OF INVENTION: Activity (5HTSA), Nucleic Acids Coding for These
 ; TITLE OF INVENTION: Polypeptides and Uses Thereof
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Rhone-Poulenc Rorer Inc.
 ; STREET: 500 Arcola Road, 3C43
 ; CITY: Collegeville
 ; STATE: PA
 ; COUNTRY: USA
 ; ZIP: 19426-0107
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/356,405
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: WO PCT/FR93/00650
 ; FILING DATE: 29-JUN-1992
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: FR 92/08081
 ; FILING DATE: 01-JUL-1992
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Smith, Julie K
 ; REGISTRATION NUMBER: 38,619
 ; REFERENCE/DOCKET NUMBER: EX92004-US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (610)454-3839
 ; TELEFAX: (610)454-3808
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 357 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-356-405-2

Alignment Scores: 2,17e-63 Length: 357
 Pred. No.: 1022.50 Matches: 219
 Score:


```
185 CTAGTGTCTGTCGTCGTCACCTTCTCTGGACCTGCTGCTCCGGTACACCTCCG 244
186 |||||
187 LeuGlyPheLeuValAlaAlaThrPheAlaThrProAsnLeuLeuValAlaThrLeuLeu 66
188 |||||
245 CGGTGCTGCTTCCACCGGTGCGCATAACTTGGTGGCTCGACGGCGCTCTCGGAC 304
246 |||||
247 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86
248 |||||
305 GAACCTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTCTCGACCGGGCGA 364
306 |||||
307 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
308 |||||
365 CGTGGCTGTCGGCGGAGCTGTGCCACCTGTGGATCTCTTCGACCGCGGAGCTGT 424
366 |||||
367 ArgTrpGlnLeuGlyArgLeuGlnLeuTrpLeuAlaCysAspVal----- 122
368 |||||
425 CTGTGCTCCCGCGGCTCGGGAACGTGGCGCATCGCGCTGGCGGCGGCGGCGCC 484
426 |||||
427 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTrpSer 142
428 |||||
485 ATCACAGCGACCTGACACACGCTGGCGACCGCGACCGCGCTCTGCTCATGATC 544
486 |||||
487 IleThrArgHisMetGluThrLeuArgThrArgLysCysValSerAsnValMetIle 162
488 |||||
545 GCGTCTGCGCGGCTCGGCTCGGCTCATCGCTCGCGCGCTGCTCTTTGGCGGCGC 604
546 |||||
547 AlaLeuThrTrpAlaLeu***ThrValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
548 |||||
605 GAGGTGTCGACGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 664
606 |||||
607 GluThrTrpSerGluGlySerGluGlyCysGlnValSerArgGluProSerTrpAlaVal 202
608 |||||
665 TTCTCCACCGCGGCTTCCACCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 724
666 |||||
667 PheSerThrValGlyAlaPheTrpLeuProLeuCysValValLeuPheValTrpLys 222
668 |||||
725 ATCTACGAGGGCGCAAGTTCGTTGCGC---CGCGCGGAGAGCTGCTGCGCGTGT 781
726 |||||
727 IleTrpLysAlaThrLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242
728 |||||
782 CCGGCGACATCATGCTGAGTCCAGGTCCTAAGGTAAGGACGACCTGATGAGCTGAGTGTG 841
783 |||||
784 SerGluAlaValGlu-----ValLysAspSerAlaGlnProGlnMetVal 258
785 |||||
842 TTCACGGCACATGCAAGACGAGTCTTCCAGGTGAGCGGGAGCTCTCGCGGGAG 901
843 |||||
844 PheThrValArg---HisAlaThrValThrPheGlnProGluGlyAspThrCysArgGlu 277
845 |||||
902 CAGAGGAGAGGAGCAGCAGCAGTATGTTGGGAATCTGATGCGGTGTTGCTGCTGTC 961
903 |||||
904 GlnLysGluGlnArgProAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297
905 |||||
962 TGGATCCCTCTTCTCTACGGAATCATATCAGCCCACTCTGTGCTCGACGCTGCGCCGCC 1021
963 |||||
964 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317
965 |||||
1022 ATCTGGAAAGCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081
1023 |||||
1024 IleTrpLysSerIlePheLeuTrpLeuGlyTrpSerAsnSerPhePheAsnProLeuIle 337
1025 |||||
1082 TACACAGCTTTTAAAGAACTACACAAATGCTTCAAGAGCTCTTTACTAAGCAG 1138
1083 |||||
1084 TyrThrAlaPheAsnLysAsnTrpAsnSerAlaPheLysAsnProPheSerArgGln 356
1085 |||||

RESULT 4
US-08-031-538-4
; Sequence 4, Application US/08031538
; Patent No. 5968817
; GENERAL INFORMATION:
; APPLICANT: Sutcliffe, J Gregor
; APPLICANT: Erlander, Mark G
; APPLICANT: Lovenberg, Timothy W
```



```
QY 500 CAGCACACGCTGCGCACCGCGCGCTGCTGCTCATGCTCGCGCTCGCGCGGTG 559
Db 148 GLeuTyrThrLeuArgAlaArgLysArgValSerAsnValMetIleLeuLeuThrTrpAla 167
QY 560 CGTTCGCGCTCATGCGCTCGCGCGCTGCTGCTTTGGCGCGGCGAGGTGTGCGAGCGT 619
Db 168 LeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGlyGluThrTyrSerGlu 187
QY 620 CGGCTCCAGCGCTCCAGGTGAGCGCGGACCTTCTATGCGCGCTTCTCCACCGCGGC 679
Db 188 LeuSerGluGluCysGlnValSerArgGluProSerTyrThrValPheSerThrValGly 207
QY 680 GCCTTCACCTGCGCTGCGCTGCGCTGCTGCTTCTACCGAAGATCTACGAGCGCGCC 739
Db 208 AlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLysIleTyrLysAlaAla 227
QY 740 AAGTTTCGTTTCGCG---CGCGCGCGAGAGCTGTGCTGCGCGTTCGCGCGCACCATGAG 796
Db 228 LysPheArgMetGlySerArgLysThrAsnSerValSerProIleProGluAlaValGlu 247
QY 797 GTGAGGTCCAAAGTAAAGGAGCACCTGATGAGGCTGAGTGGTGTTCACGGCACATGC 856
Db 248 -----ValLysAspAlaSerGlnHisProGlnMetValPheThrValArg--- 262
QY 857 AAAGCAACGCTGTCTTCCAGTGAGCGGGGACTCTGCGGGGAGCAGAGGAGGCGGA 916
Db 263 HisAlaThrValThrPheGlnThrGluGlyAspThrTrpArgGluGlnLysGluGlnArg 282
QY 917 GCAGCCATGATGGGAATCTGATTGGCGTGTGTGTGCTGTGCTGATCCCTTCTTTC 976
Db 283 AlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCysTyrPheProPhePhe 302
QY 977 CTGACGGAATCATGACCGCACTGTGCTGCGAGCTGCGCGCTGCGCGCGCTGCGAAGCATA 1036
Db 303 ValThrGluLeuIleSerProLeuCysSerTrpAspIleProAlaLeuTrpLysSerIle 322
QY 1037 TTTCTGCGCTGGCTACTCTCAATCTTCTTCAACCCCTGATTACACAGCTTTTAAC 1096
Db 323 PheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyrThrAlaPheAsn 342
QY 1097 AAGAACTACAACTGCTTCAAGAGCTCTTTACTAAGCAGAGA 1141
Db 343 ArgSerTyrSerSerAlaPheLysValPhePheSerLysGlnGln 357
```

RESULT 5

```
US-08-461-812-4
; Sequence 4, Application US/08461812
; Patent No. 5935925
; GENERAL INFORMATION:
; APPLICANT: Weinshank, Richard L.
; APPLICANT: Branchek, Theresa
; APPLICANT: Hartig, Paul R.
; TITLE OF INVENTION: DNA ENCODING HUMAN 5-HT1D RECEPTORS AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,812
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
```

```
;
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36536-B/JPW/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 2122780400
; TELEFAX: 2123910525
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-812-4
```

Alignment Scores:

```
Pred. No.: 1.43e-23 Length: 390
Score: 448.00 Matches: 127
Percent Similarity: 48.35% Conservative: 63
Best Local Similarity: 32.32% Mismatches: 147
Query Match: 20.80% Indels: 56
DB: 2 Gaps: 13
```

US-09-976-782-15 (1-1152) x US-08-461-812-4 (1-390)

```
QY 59 CCGGAGACAGCAGCGGAGCGGAGCCCAAGCCGAGAGGATATCTGGTTTCACCCCG 118
Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
QY 119 AGC-----GGCGCCGCTCTGCGCGGCGCGA 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
QY 143 GGGCGCGCTTCTGCTTTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
QY 203 GCCACTTTCCTGTGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
QY 263 CGGTGCGCGCATAACTTGTGGCTCGACGCGCTGCTGCGAGCAACTAGTGGGAGCGGTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
QY 323 CGCATGCGCAGCGCTGCGGAGTGTGCTGACCGCGGCGGCGCTGCTGCTGCTGCTGCTG 382
Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
QY 383 AGCTGTGCGCACGTGTGATCTCTTCCAGCGCGGAGCTGCTGCTGCTGCTGCTGCTGCTG 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
QY 443 CTGCGGAACTGCGCGCATGCGCTGCGCGCGGAGCGGCGCATACACGCGACCTGCGAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
QY 503 CACAGCTGCGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
QY 563 TCGGCGCTATCGCCCTCGCGCGCTGCTTTCGCGCGGCGGAGGTGTGCGAGCTCGG 622
Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGluGlu 195
QY 623 CTCAGCGCTGCGAGGTGAGCGCGGAACCC---TCCTATCCGCGCTTCTCCACCGCGCGC 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 739
Db 216 AlaPheTyrPheProThrLeuLeuAlaLeuTyrGlyArgIleTyrValGluAla 235
QY 740 AAGTTTCGTTTC-----GGCGCGCGC-----CGGAGAGCTGTG 772
```

Db 236 ArgSerArgLeuLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTTGCGGCGCCACCATCGAGTCCAAAGCTAAG 814
Db 256 LeuLeuThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCTGATGAGCT---GAAGTGGTGTTCACGGCACATTCGAAAGCAACGGGTGCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGACTCTCGCGGAGCAGAAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuLysLysLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CGAGCAGCATGATGGTGGGAATCTGATTGGCGTGTGTGCTGTGCTGGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
QY 974 TTCCTGACGGAACTCATCAGCCCACTCTGT-----GCCTGCAGCCTGCCCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY 1028 AAAAGCATATTTCTGTGGCTTGGCTACTCCATTTCTTCTTCAACCCCTGATTTACACA 1087
Db 351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
QY 1088 GCTTTTAAACAAGAACTACAAATGCTTCAAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

RESULT 6

US-08-157-185-15
; Sequence 15, Application US/08157185
; Patent No. 5985585
; GENERAL INFORMATION:
; APPLICANT: Bard A. Jonathan
; APPLICANT: Branchek A. Theresa
; APPLICANT: Weinshank L. Richard
; TITLE OF INVENTION: DNA ENCODING A HUMAN SEROTONIN
; TITLE OF INVENTION: RECEPTOR (5-HT4B) AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,185
; FILING DATE: 15-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White P., John
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 41908-A-PCT-US/JPW/MAT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-157-185-15

Alignment Scores:
Pred. No.: 1,43e-23 Length: 390
Score: 448.00 Matches: 127
Percent Similarity: 48.35% Conservative: 63
Best Local Similarity: 32.38% Mismatches: 147
Query Match: 20.80% Indels: 56
DB: 2 Gaps: 13
US-09-976-782-15 (1-1152) x US-08-157-185-15 (1-390)
QY 59 CCGAGAGACAGCAGCGGACCCCGGGACCCAGAGGAGGATGACTCGTTCGACCCCG 118
Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
QY 119 AGC-----GGCGCGCTCTCTCGCGGCCGA 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
QY 143 GGGCGCCCTTCTCTGTCTTACGGTCTGTGGTGGTGGTCTCTAGTGTCTGTCTGCT 202
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
QY 203 GCCACTTCTCTGTGAACCTGCTGCTCGGTTCGGTTCACATCCCGCGGGTCCGTCTCCAC 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
QY 263 CGCGTGGCGCATAACTTGTGGCTTCGACGCGCGCTCGACGAACTAGTGGCAGCGCTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
QY 323 GCGATGCCACCGAGCTGGCGAGTGCTGACGCGGGCGACGCGGTCTGTGGCGCGG 382
Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
QY 383 AGCTGTGCGCAGCTGTGATCTCTTCGACGCGCGGAGCGTGTCTGTGCTGCCCGCGCGC 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
QY 443 CTCGGGAACGTGGCGCCATCGCCCTGGCGCGCGACGGGCCCATCACAGCGACCTGCAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
QY 503 CACAGCTGCGCACCGCAGCGCGCTGTGTTCATGATCGCGTCCCGCGGGTGGCG 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
QY 563 TCGGCGCTCATCGCCCTCGCGCGCTGCTTTTGGCGGGCGGAGGTGTGCCAGCTCGG 622
Db 177 SerIleSerIleSerLeuProPro-----PhePheTrpArgGlnAlaLysAlaGluGlu 195
QY 623 CTCAGCGCTGCGCAGTGAGCGCGGAACCC---TCCTATCGCGCTTCTCCACCCCGCGC 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTCGCGTGGCGTGTGTGTTCACCGGAAGATCTACGAGGGCGGC 739
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
QY 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGCTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTTGCGGCGCACCATCGAGTGAGGTCCAAAGTAAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCTGATGAGCT---GAAAGTGGTTCACGGCACATTCGAAAGCAACGGGTGCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGGACTCTCGCGGAGCAGAAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310

Qy	563	TCGCGGCTATATCGCCCTCGCGCCGCTCTTTGGCCGCGGCGAGGTGTGGCAGCGCTCGG	622
Db	177	SerIleSerIleSerLeuProProPhePheTrpArgGlnAlaLysAlaGluGluGlu	195
Qy	623	CTCCAGCGCTGCAGAGTGAGCCGGGAACCC---TCCTATGCGCGCTTCTCCACCGCGGC	679
Db	196	ValSerGluCysValValAenThrAspHisIleLeuThrValTyrSerIthrValGly	215
Qy	680	GCCTTCACACTGCGCTGGCGTGGTCCGCTTCTACCGAAGATCTACGAGCGCGCC	739
Db	216	AlaPheTyrPheProThrLeuLeuLeuAlaLeuTyrGlyArgIleTyrValGluAla	235
Qy	740	AAGTTTCGTTTC-----GGCGCGCCG---CGGAGAGCTGTG	772
Db	236	ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln	255
Qy	773	CTG-----CCGTTCCGCGCCACCATGCAGGTGAGGTCCAAAGTAAAG	814
Db	256	LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro	275
Qy	815	GAGACACTGATGAGGCT---GAAGTGGTGTTCACGGCACATTGCCAAGACACGGTGTCC	871
Db	276	AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer	295
Qy	872	TTCAGGTGAGCGGGACTCTCGCGGAGCAGAG---GAGAGG	913
Db	296	-----AspAlaLeuLeuGlnLysLysLeuMetAlaAlaArgGluArg	310
Qy	914	CGAGCAGCCATGATGGGGAATCTGATGGCGTGTGTGCTGCTGATCCCGTTC	973
Db	311	LysAlaThrLysThrLeuGlyIleLeuLeuGlyAlaPheIleValCysTrpLeuProPhe	330
Qy	974	TTCTCTGACGAACTCATCAGCCCACTCTGT-----GCCTGCAGCTGCCCGCCATCTGG	1027
Db	331	PheIleLeuSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle	350
Qy	1028	AAAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTTTTCACACCCCGCTGATTTACACA	1087
Db	351	PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr	370
Qy	1088	GCITTTAAAGAACTACAAATGCCTTCAAGAGCTC	1126
Db	371	MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu	383

RESULT 9

US-09-332-837-15

; Sequence 15, Application US/09332837

; Patent No. 8432855

; GENERAL INFORMATION:

; APPLICANT: Bard, Jonathan

; APPLICANT: Branchek, Theresa

; APPLICANT: Weinschank, Richard

; TITLE OF INVENTION: Methods of Obtaining Pharmaceutical Compositions

; FILE REFERENCE: 41908-AA-PCT-US

; CURRENT APPLICATION NUMBER: US/09/332,837

; CURRENT FILING DATE: 1999-06-14

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 15

; LENGTH: 390

; TYPE: PRT

; ORGANISM: Homo Sapiens

US-09-332-837-15

Alignment Scores:		
Pred. No.:	1,43e-23	Length: 390
Score:	448.00	Matches: 127
Percent Similarity:	48.35%	Conservative: 63
Best Local Similarity:	32.32%	Mismatches: 147
Query Match:	20.80%	Indels: 56
DB:	4	Gaps: 13


```
QY 623 CTCACGCGTGCAGGTGAGCGGGAACCC---TCCTATGCCGCTTCTCCACCGCGGC 679
Db 196 ValSerGluCysValAlaSerThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTGCGCTTGGCGTGGCGTGGCTTCTACCGGAAGATCTACGAGCGCGCC 739
Db 216 AlaPheTyrPheProThrLeuLeuLeuLeuAlaLeuTyrGlyArgIleTyrValGluAla 235
QY 740 AAGTTTGGTTTC-----GGCCCGCGC---CGGAGAGCTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyTyrAspArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTGGCGCCACCATGCAGCTGAGGTCCAAAGTAAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCTGATGAGGCT---GAAGTGGTGTTCACGGCACATTCGCAAGCAACGGTGTCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGGGACTCCTGGCGGGAGCAGAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CGAGCAGCCATGATGGTGGGAATCTGATTCGGGTGTTGTGCTGTGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleLeuLeuGlyAlaPheIleValCysTrpPheProPhe 330
QY 974 TTCCTGACGGAACTCATCAGCCCACTCTGT-----GCTCGCAGCCTGCCCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY 1028 AAAAGCATATTCGTGGCTGGCTACTCCAAATCTTCTCAACCCCTCATTTACACA 1087
Db 351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
QY 1088 GCTTTTAAAGAACTACAACAATGCCCTTCAAGACCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

RESULT 12
US-08-117-006-6
; Sequence 6, Application US/08117006
; Patent No. 5639652
; GENERAL INFORMATION:
; APPLICANT: Weinshank, Richard L.
; APPLICANT: Branchek, Theresa
; APPLICANT: Hartig, Paul R.
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR AND
; TITLE OF INVENTION: USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESS: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect, Version 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,006
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1795/39318
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
```

```
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: 5-HT1DB
US-08-117-006-6

Alignment Scores:
Pred. No.: 1,56e-22 Length: 390
Score: 433.00 Matches: 124
Percent Similarity: 47.84% Conservative: 64
Best Local Similarity: 31.55% Mismatches: 149
Query Match: 20.10% Indels: 56
DB: 1 Gaps: 13

US-09-976-782-15 (1-1152) x US-08-117-006-6 (1-390)
QY 59 CCCGAGACCGACGAGCGGACCGGACCCCAAGCCGAGAGGATACGCTTCGACCCCG 118
Db 10 ProProAlaProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
QY 119 AGC-----GGCGCGCTCCTCCCGCGCGA 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
QY 143 GGGCGCGCTTCTCTGCTTTCACGGTCTGTGTGTGAGCGCTGTAGTGTCTGATCGCT 202
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
QY 203 GCCATTTCTGTGGAACCTCTGCTTCCGGTCCGGTCCACCATCCCGCGGTCCGCTTCAC 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
QY 263 CGCGTCCGCATAACTTGGTGGCTTCGACGGCGCTCTCGACGAACTAGTGGCAGCGCTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAspValThrAspLeuValSerIleLeu 101
QY 323 GCGATGCCACGACGCTGGCGAGTGTGTGCGACCGGGCGACGTCGCGCTGCTGGCGCG 382
Db 102 ValIlePro-----IleSerThrMetTyrThrValThrAspArgTrpThrLeuSerGln 119
QY 383 AGCTGTGCCACGTGTGATCTCTTCGACCGCGAGCGCTCTGTGCTGCTCCCGCGCGC 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
QY 443 CTCGGGAACGTGGCGGCATCGCTTGGCGCGCGCGCGCGCATCACACGACCATCGAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
QY 503 CACAGCTGCGCACCGCGCGCGCTGTGTGTGTATCATGCTGCGCTGCGCGGTGCGCG 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
QY 563 TCGGGCTCATCGCCCTCGCGCGCTGCTCTTTCGGCGGGGAGGTGTGGACGCTCGG 622
Db 177 SerIleSerIleSerLeuProPro----PhePheTrpArgGlnAlaLysAlaGluGlu 195
QY 623 CTCGAGCGCTGCCAGGTGAGCGGGAACCC---TCCTATGCCGCTTCTCCACCGCGCG 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTGCGCTTGGCGTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 739
Db 216 AlaPheTyrPheProThrLeuLeuLeuLeuAlaLeuTyrGlyArgIleTyrValGluAla 235
```



```
QY 740 AAGTTTCGTTTC-----GGCGGCGC---CGGAGAGCTGTG 772
   : : : : :
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
   : : : : :
QY 773 CTG-----CGTTGCGCGCCACCATCGAGGTGAGGTCCAAAGTTAAAG 814
   : : : : :
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
   : : : : :
QY 815 GAAGCACCTGATGAGGCT--GAAGTGGTGTTCACGCGCACATTGCAAGCAACGGGTGTC 871
   : : : : :
Db 276 AspValProSerGluSerGlySerProValThrValAsnGlnValLysValArgValSer 295
   : : : : :
QY 872 TTCAGGTGAGCGGGGACTCCTGCGGAGCAGAGAAG-----GAGAGG 913
   : : : : :
Db 296 -----AspAlaLeuLeuGluLysLysLeuMetAlaAlaArgGluArg 310
   : : : : :
QY 914 CGAGCAGCATGATGTTGGGAATCTGATGGGTGTTTGTGCTGTGCTGATCCCTTC 973
   : : : : :
Db 311 LysAlaThrLysThrLeuGlyIleLeuLeuGlyAlaPheIleValCysTrpLeuProPhe 330
   : : : : :
QY 974 TTCCTCAGCGAACTCATCAGCCACACTGT-----GCCTGACAGCCTGCCCCCATCTGG 1027
   : : : : :
Db 331 PheIleLysSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
   : : : : :
QY 1028 AAAAGCATATTTCTGTGGCTGGCTACTCCAAATTTCTTCAACCCCTGATTTACACA 1087
   : : : : :
Db 351 PheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
   : : : : :
QY 1088 GCTTTTAAACAGAACTACACAAATGCTTCAAGAGCCTC 1126
   : : : : :
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383
```

RESULT 13

```
US-08-216-594-6
; Sequence 6, Application US/08216594
; Patent No. 5652113
; GENERAL INFORMATION:
; APPLICANT: Weinshank, Richard L.
; APPLICANT: Branchek, Theresa
; APPLICANT: Hartig, Paul R.
; TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1F RECEPTOR
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,594
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 1795/39318
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
```

```
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: 5-HT1DB
; US-08-216-594-6
```

Alignment Scores:

Pred. No.:	1,56e-22	Length:	390
Score:	433.00	Matches:	124
Percent Similarity:	47.84%	Conservative:	64
Best Local Similarity:	31.55%	Mismatches:	149
Query Match:	20.10%	Indels:	56
DB:	1	Gaps:	13

US-09-976-782-15 (1-1152) x US-08-216-594-6 (1-390)

```
QY 59 CCGGAGACACGAGCGGACCGGACCCCAAGCCCGAGAGGAGGATCTCGTTCGACCCCG 118
   : : : : :
Db 10 ProProAlaProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerAlaPro 29
   : : : : :
QY 119 AGC-----GGCGCCGCTCTGCGGCGCGA 142
   : : : : :
Db 30 SerGlnAspCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
   : : : : :
QY 143 GGGCGCCGCTCTCTCTGCTTACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
   : : : : :
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
   : : : : :
QY 203 GCCACTTCTCTGGAACCTGCTGCTCGGTCCGATCCCGGGTCCGTCGCTTCCAC 262
   : : : : :
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
   : : : : :
QY 263 CCGGTGCCCATAACTTGTGTCGCTGCGAGCGGCTCTCGAGAGAACTAGTAGGACGCTG 322
   : : : : :
Db 82 ThrProAlaAsnTyrIleIleAlaSerLeuAspValThrAspLeuLeuValSerIleLeu 101
   : : : : :
QY 323 GCGATGCCACCGAGCTGCGAGTGAGCTGTCGCGGCGGCGAGCTGCTGCTGCTGCTGCTG 382
   : : : : :
Db 102 ValIlePro-----IleSerThrMetTyrThrValThrAspArgTrpThrLeuSerGln 119
   : : : : :
QY 383 AGCCTGTGCCACGTGTGATCTCTTCGACGCGGAGCCTGTCTGTGCTGCCCGCGCG 442
   : : : : :
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
   : : : : :
QY 443 CTCGGGAGCTGGCGGCATCGCCCTGGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 502
   : : : : :
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
   : : : : :
QY 503 CACAGCTGCGCACCGCGAGCGCTGTTGCTCATGATCGCTGCTGCTGCTGCTGCTGCTG 562
   : : : : :
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
   : : : : :
QY 563 TCGCGCTCATCGCCCTCGCGCTGCTGCTTTTGGCGGGGCGGAGGTGTGCGACGCTCG 622
   : : : : :
Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
   : : : : :
QY 623 CTCGAGCGCTGCGAGTGAGCGGGAACCC---TCCTATGCGGCTCTTCCGACCCGCG 679
   : : : : :
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
   : : : : :
QY 680 GCTTTCCACCTGCGCTTGGCTGCTGCTTGTCTACCGGAAGATCTACGAGCGGCG 739
   : : : : :
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
   : : : : :
QY 740 AAGTTTCGTTTC-----GGCGGCGC---CGGAGAGCTGTG 772
   : : : : :
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
   : : : : :
QY 773 CTG-----CCGTTGCGGCGCACCATGTCAGGTGAGGTCCAAAGTTAAAG 814
   : : : : :
```


Db 296 -----AspAlaLeuLeuGluLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CGAGCAGCATCATGGTGGGAATTCGATTGGCGCTGTTGTGCTGCTGGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
QY 974 TTCCTGAGCGAACTCATGAGCCCACTCTGT-----GCCTGCAGCGCTGCCCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY 1028 AAAAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTTCTTCAACCCCTGATTACACA 1087
Db 351 PheAspPheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
QY 1088 GCTTTTACAGAACTACAACTGCCTTCAAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

RESULT 15
US-08-370-542-6
Sequence 6, Application US/08370542
Patent No. 5476782
GENERAL INFORMATION:
APPLICANT: Weinshank, Richard L.
APPLICANT: Brancheck, Theresa
APPLICANT: Hartig, Paul R.
TITLE OF INVENTION: DNA ENCODING A HUMAN 5-HT1E RECEPTOR AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESS: Cooper & Dunham
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/370,542
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/194,113
FILING DATE:
APPLICATION NUMBER: US/07/803,626
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 1795/39317
TELEPHONE: 212-977-9550
TELEFAX: 212-664-0525
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 398 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-370-542-6

Alignment Scores: 1.57e-22 Length: 398
Pred. No.: 433.00 Matches: 124
Score:

Percent Similarity: 47.84% Conservative: 64
Best Local Similarity: 31.55% Mismatches: 149
Query Match: 20.10% Indels: 56
DB: 1 Gaps: 13
US-09-976-782-15 (1-1152) x US-08-370-542-6 (1-398)
QY 59 CCGAGAGACCAGACGGGACCCGGGACCCAGCCGAGAGGATCTCGTTTCGACCCCG 118
Db 10 ProProAlaProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
QY 119 AGC-----GGCGCCGCTCCGCGGCCGA 142
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
QY 143 GGGCGCCCTTCTCTGCTTTCACGGCTGCTGGTGGTGACGCTGCTAGTGTGCTGATCGCT 202
Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
QY 203 GCCACTTCTCTGGAACTGCTGCTTCCGCTCACCATCCCGCGGGTCCGCTTCCAC 262
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
QY 263 CCGGTGCGCATAACTGTTGGCTTCGACGGCGCTCTCGAGCAACTAGTGTGCGAGCGCTG 322
Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAspValThrAspLeuLeuValSerIleLeu 101
QY 323 CGATGCCACCGAGCTGCGAGTGTGTGCGCGGCGGCGCTGCTGGCGCG 382
Db 102 ValIlePro-----IleSerThrMetTyrThrValThrAspArgTrpThrLeuSerGln 119
QY 383 AGCCTGTGCGACGTGTGGATCTCTTCGACGCGGAGCCTCTGTGTGCTGCCCGCGCG 442
Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
QY 443 CTCGGGAACGTGGCGGCATCGCCTGGCGCGGACGGGCGCATCACAGCGCACCTGCGAG 502
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
QY 503 CACAGCTGCGCACCGCGCGCGCTGTTGCTCATGATCGCGCTCGCCCGGGTGGCG 562
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
QY 563 TCGGGGCTCATCGCCCTCGCGCGCTGCTTTTGGCGGGGAGGTGTGCGACGCTCGG 622
Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
QY 623 CTCAGCGCTGCCAGTGTAGCGGGAACCC-----TCGTATGCGCGCTTCTCCACCGCGCG 679
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
QY 680 GCCTTCCACCTCGCGCTTGGCGTGTGCTTGTCTACCGGAAGATCTACGAGGCGCG 739
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
QY 740 AAGTTTCGTTTC-----GGCGCGCGC---CGGAGAGCTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTTGCGGCGCACCATCGAGTGTGAGTGTCCAGGTAAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCTGATGAGGT---GAAGTGGTGTTCAGCGCACATTGCAAAACAACGGTGTCC 871
Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGAGCTCTCGCGGGAGCAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CGAGCAGCATCATGGTGGGAATTCGATTGGCGCTGTTGTGCTGCTGGATCCCTTC 973
Db 124

Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTipLeuProPhe 330
Qy 974 TTCTGTGACGGAACATCATCAGCCCACTCTGT-----GCCTGCAGCCTGCCCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
Qy 1028 AAAAGCATATTTCTGTGGCTTGCTACTCCAAATCTTTCTTCAACCCCTGATTACACA 1087
Db 351 PheAspPheThrTrpLeuGlyTyLeuAsnSerLeuIleAsnProIleIleTyThr 370
Qy 1088 GCTTTTAACAAGAACTACAACAATGCCTTCAAGAGCCTC 1126
Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383

Search completed: August 31, 2004, 20:24:14
Job time : 47 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 19:59:22 ; Search time 28.5 Seconds

(without alignments)
4209.461 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcagtgaggccgtagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 283362

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model p/US09976782/runat_31082004_115950_10303/app_query.fasta_1.1351
-Q=/cgn2_1/USPTO_spool_p/US09976782/runat_31082004_115950_10303/app_query.fasta_1.1351
-DB=SwissProt_42 -OFMT=fastan -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09976782 @CGN 1 1 35 @runat_31082004_115950_10303 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THRADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB ID	Description
1	1453.5	67.5	370	1	SH5B_RAT	P35365 rattus norv
2	1452.5	67.4	370	1	SH5B_MOUSE	P31387 mus musculu
3	1063.5	49.4	357	1	SH5A_HUMAN	P47898 homo sapien
4	1022.5	47.5	357	1	SH5A_MOUSE	P30966 mus musculu
5	1014.5	47.1	357	1	SH5A_RAT	P35364 rattus norv
6	448	20.8	390	1	SH1B_HUMAN	P28222 homo sapien
7	448	20.8	390	1	SH1B_PANTR	P60020 pan troglod
8	447	20.8	386	1	SH1B_SAEH	P56496 spalax leuc
9	444.5	20.6	389	1	SH1B_CAVPO	O08892 cavia porce
10	444	20.6	390	1	SH1B_RABIT	P49144 oryctolagus
11	441.5	20.5	564	1	SH1I_DROME	P20905 drosophila
12	441	20.5	377	1	SH1D_CANFA	P11614 canis fami
13	440	20.4	386	1	SH1B_CRIGR	P46636 cricetulus
14	438	20.3	386	1	SH1B_MOUSE	P28334 mus musculu
15	437.5	20.3	376	1	SH1D_CAVPO	Q60484 cavia porce
16	437	20.3	386	1	SH1B_RAT	P28564 rattus norv
17	435.5	20.2	416	1	SH1B_FUGRU	O42384 fugu rubrip
18	430	20.0	377	1	SH1D_HUMAN	P28221 homo sapien

19	429.5	19.9	446	1	SH7_BOMMO	Q17239 bombyx mori
20	428	19.9	422	1	SH1A_HUMAN	P08908 homo sapien
21	428	19.9	422	1	SH1A_PANTR	Q9n298 pan troglod
22	425.5	19.8	379	1	SH1D_FUGRU	P79748 fugu rubrip
23	425	19.7	421	1	SH1A_MOUSE	O64264 mus musculu
24	421	19.5	422	1	SH1A_RAT	P19327 rattus norv
25	419.5	19.5	374	1	SH1D_MOUSE	Q61224 mus musculu
26	418	19.4	466	1	SH7_HUMAN	Q25190 heliothis v
27	417	19.4	479	1	SH7_HUMAN	Q25414 lymnaea sta
28	416.5	19.3	509	1	SH7_LYMS	P34969 homo sapien
29	416	19.3	374	1	SH1D_RAT	P28565 rattus norv
30	416	19.3	377	1	SH1D_RABIT	P49145 oryctolagus
31	416	19.3	446	1	SH7_CAVPO	P50407 cavia porce
32	413.5	19.2	366	1	SH1F_CAVPO	O08890 cavia porce
33	412.5	19.2	366	1	SH1F_HUMAN	P30939 homo sapien
34	412.5	19.2	388	1	SH1B_DIDMA	P35404 didelphis m
35	411.5	19.1	365	1	SH1F_PANTR	Q9n2d9 pan troglod
36	411.5	19.1	366	1	SH1F_MOUSE	Q02284 mus musculu
37	409.5	19.0	379	1	GRE2_BALAM	Q93127 balanus amp
38	408	18.9	363	1	SH1E_PANTR	Q9n2b6 pan troglod
39	408	18.9	365	1	SH1E_HUMAN	P28566 homo sapien
40	408	18.9	448	1	SH7_RAT	P32305 rattus norv
41	408	18.9	484	1	OAR1_LOCMI	Q25321 locusta mig
42	408	18.9	484	1	OAR2_LOCMI	Q25322 locusta mig
43	406.5	18.9	366	1	SH1F_RAT	P30940 rattus norv
44	406.5	18.9	423	1	SH1A_FUGRU	O42385 fugu rubrip
45	403	18.7	448	1	SH7_MOUSE	P32304 mus musculu

ALIGNMENTS

RESULT 1

SH5B_RAT	SH5B_RAT	STANDARD;	PRT;	370 AA.
ID	P35365;			
AC	01-JUN-1994 (Rel. 29, Created)			
DT	01-JUN-1994 (Rel. 29, Last sequence update)			
DT	01-OCT-1996 (Rel. 34, Last annotation update)			
DE	5-hydroxytryptamine 5B receptor (5-HT-5B) (Serotonin receptor) (MR22).			
GN	HTS5B OR HTS5B.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]_TaxID=10116;			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;			
RX	MEDLINE=93234515; PubMed=7682702;			
RA	Erlander M.G., Lovenberg T.W., Baron B.M., de Lecea L.,			
RA	Danielson P.E., Racke M., Slone A.L., Siegel B.W., Foye P.E.,			
RA	Cannon K., Burns J.E., Sutcliffe G.J.,			
RT	"Two members of a distinct subfamily of 5-hydroxytryptamine receptors			
RT	differentially expressed in rat brain."			
RL	Proc. Natl. Acad. Sci. U.S.A. 90:3452-3456(1993).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=94039744; PubMed=8224165;			
RA	Wisden W., Parker E.M., Mahle C.D., Grisel D.A., Nowak H.P.,			
RA	Yocca F.D., Felder C.C., Seeburg P.H., Voigt M.M.;			
RT	"Cloning and characterization of the rat 5-HT5B receptor. Evidence			
RT	that the 5-HT5B receptor couples to a G protein in mammalian cell			
RT	membranes."			
RL	FEBS Lett. 333:25-31(1993).			
CC	!- FUNCTION: This is one of the several different receptors for 5-			
CC	hydroxytryptamine (serotonin), a biogenic hormone that functions			
CC	as a neurotransmitter, a hormone, and a mitogen. The activity of			
CC	this receptor is mediated by G proteins. Probably involved in			
CC	anxiety and depression.			
CC	!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	!- TISSUE SPECIFICITY: Brain; in the CA1 region of hippocampus, the			
CC	medial habenula, and raphe nuclei.			
CC	!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.			
CC	STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L10073; AAA40616.1; --
 CC DR PIR; S38744; S38744;
 CC DR InterPro; IPR00276; GPCR_Rhodopsn.
 CC DR Pfam; PF00001; 7tm_1; 1.
 CC DR PRINTS; PR00237; GPCR_RHODOPSIN.
 CC DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 CC DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 CC KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC KW Multigene family.
 CC -----

DR EMBL; L10073; AAA40616.1; --
 DR PIR; S38744; S38744;
 DR InterPro; IPR00276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCR_RHODOPSIN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 53 75 1 (POTENTIAL).
 FT DOMAIN 76 90 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 91 111 2 (POTENTIAL).
 FT DOMAIN 112 128 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 129 150 3 (POTENTIAL).
 FT DOMAIN 151 171 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 172 194 4 (POTENTIAL).
 FT DOMAIN 195 211 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 212 232 5 (POTENTIAL).
 FT DOMAIN 233 295 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 296 316 6 (POTENTIAL).
 FT DOMAIN 317 333 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 334 354 7 (POTENTIAL).
 FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 5 5 N-LINKED (GLCNAC...) (POTENTIAL).
 FT DISULFID 127 205 BY SIMILARITY.
 SQ SEQUENCE 370 AA; 41122 MW; 8ECSF79BFD647E5 CRC64;

Alignment Scores:

Pred. No.: 3,25e-67 Length: 370
 Score: 1453.50 Matches: 296
 Percent Similarity: 83.38% Conservative: 20
 Best Local Similarity: 78.10% Mismatches: 54
 Query Match: 67.48% Indels: 9
 DB: 1 Gaps: 3

US-09-976-782-15 (1-1152) x 5HSB_RAT (1-370)

QY 5 ATGGAGCGCTAGCTTTCAGTGGCCACCGCGCGGTTCCTTCCCTGGGACCCGAG 64
 Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20
 QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCCGAGAGGATACCTCGGTTTCGACCCCGAGCGG 124
 Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
 QY 125 GCGTCTCGCGGCGCGAGGCGCGCTTCTCTGCTTTCACGCTCTCGTGTGAGCGTG 184
 Db 39 LeuLeuLeuSerGlyArgGluProPheSerAlaPheThrValLeuValValThrLeu 58
 QY 185 CTAGTCTGCTGATCGCTGCGCTTTCCTGTGGAACCTGCTGGTTCGGTTCACATCCCG 244
 Db 59 LeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuValValThrLeuLeu 78
 QY 245 CGGTCTCGCTTTCACCGCGTTCGCGCATACCTCGTGGCTTCGACGCGCTCTCGGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAACATAGTGGAAGCTGCGCGATGCCACCGAGCTCGCGAGTGAAGTGTGACCGGCGA 364
 Db 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
 QY 365 CQTGCGCTCTCGCGGCGAGCTGTGCGACGCTGTGATCTCTCTGACCGCGAGCGCTGT 424

Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
 QY 425 CTGTGCTGCCCCCGCGGCTCGGGAACGTGGCGGCATCGCCCTGGGCGCGAGCGGGGCC 484
 Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
 QY 485 ATCACAGGCACTGCGAGCACGCTCGGACCCGCGAGCGCGCTCGTTCCTCATGTC 544
 Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgArgAlaSerAlaLeuMetIle 175
 QY 545 GCGCTCGCGCGGCTCGGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGCC 604
 Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaPheLeuLeuPheGlyTrpGly 195
 QY 605 GAGTGTGCGACGCTCGGCTCCAGCGTGCAGGTGAGCGCGGAACCTCTCATGCGCGCC 664
 Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
 QY 665 TTCTCCACCCCGCGGCTTCCACCTCGCGCTTGGCGTGGTGGCGTGTGCTACCGGAG 724
 Db 216 PheSerThrCysGlyAlaPheTyrValProLeuAlaValValLeuPheValTyrTrpLys 235
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGCTGCGCTTCGCG 784
 Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255
 QY 785 GCCACCATGCGAGTGTGAGTCCCAAGGTAAAGAGACCTGATGAGGCTGAAGTGTTC 844
 Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271
 QY 845 ACCGCACTGTCAAGCAAGCGGTCTCTCCAGTGTGAGCGGCGGACTCTCGCGGAGCAG 904
 Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGCAGAGCGGACGCGCATGATGTGGGAATTCGTATGTCGCTGTGTGCTGTGCTGG 964
 Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTTCTCTGAGCAACTCATCAGCCCACTCTGTGCTGCGCTGCGCCCCCATC 1024
 Db 312 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProIle 331
 QY 1025 TGGAAAAGCATATTCTGTGGCTTGGCTACTCTCAATTCCTTCTTCAACCCCTGATTAC 1084
 Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 QY 1085 ACAGCTTTTAAAGAACTACAAATGCTTCAAGAGCTCTTTTACTAAGCAGAGA 1141
 Db 352 ThrAlaPheAsnLysAsnTyrAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 2

5HSB_MOUSE STANDARD; PRT; 370 AA.
 ID 5HSB_MOUSE
 AC P31387;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE 5-hydroxytryptamine 5B receptor (5-HT-5B) (Serotonin receptor).
 GN HTR5B OR 5HT5B.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93196607; PubMed=8450829;
 RA Matthes H., Boschert U., Amlaiky N., Grailhe R., Plassat J.-L.,
 RA Muscatelli F., Mattei M.-G., Hen R.;
 RT "Mouse 5-hydroxytryptamine5A and 5-hydroxytryptamine5B receptors
 RT define a new family of serotonin receptors: cloning, functional
 RT expression, and chromosomal localization.";
 RL Mol. Pharmacol. 43:313-319(1993).

CC -i- FUNCTION: This is one of the several different receptors for
CC 5-hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins. Probably involved in
CC anxiety and depression.
CC -i- SUBCELLULAR LOCATION: Integral membrane protein.
CC -i- TISSUE SPECIFICITY: Expressed predominantly in the central nervous
CC system; in the hippocampus, habenula, and the dorsal raphe.
CC -i- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X69867; CAA49501.1; --
CC PIR: I48231; I48231.
CC MGD: MGI:96284; Htr5b.
CC DR InterPro: IPR000276; GPCR_Rhodopsn.
CC DR Pfam: PF00001; 7tm 1; 1.
CC DR PRINTS: PR00237; GPCRHHODOPS.
CC DR PROSITE: PS00237; G-PROTEIN RECEPTOR F1_1; 1.
CC DR PROSITE: PS00262; G-PROTEIN RECEPTOR F1_2; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein;
CC Multigene family.
CC FT DOMAIN 1 52 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 53 75 1 (POTENTIAL).
CC FT DOMAIN 76 90 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 91 111 2 (POTENTIAL).
CC FT DOMAIN 112 128 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 129 150 3 (POTENTIAL).
CC FT DOMAIN 151 171 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 172 194 4 (POTENTIAL).
CC FT DOMAIN 195 211 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 212 232 5 (POTENTIAL).
CC FT DOMAIN 233 295 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 296 316 6 (POTENTIAL).
CC FT DOMAIN 317 333 EXTRACELLULAR (POTENTIAL).
CC FT TRANSMEM 334 354 7 (POTENTIAL).
CC FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).
CC FT CARBOHYD 5 5 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT DISULFID 127 205 BY SIMILARITY.
CC SQ SEQUENCE 370 AA; 41201 MW; 0553C62B12DAAD84 CRC64;

Alignment Scores:
Pred. No.: 3,66e-67 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.84% Conservative: 19
Best Local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 1 Gaps: 3

US-09-976-782-15 (1-1152) x 5H5B_MOUSE (1-370)

Qy 5 ATGGAGGCGCGTAGCCTTCAGTGGCCACCGCGCGGTGGCTTGCCTCGGACCCGCGAG 64
Dbb 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProProGlyProGlu 20
Qy 65 ACCAGCAGCGACCGGACCGGACCGGAGGAGGACTCGGTTCCAGCCCGGCGGC 124
Dbb 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGly 38
Qy 125 GCGGTCCTGCGCGGCGGAGGCGCGCTTCTGCTTCAGGTCCTCGTGGTGGACGTG 184
Dbb 39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValValValThrLeu 58
Qy 185 CTAGTGTGCTGATCGCTGCCACTTTCCTGTGGAACTGCTGGTCCGCTCACCATCCG 244
Dbb 59 LeuValLeuLeuAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78

Qy 245 CGGTCGCTGCTTCCACCGCGTCCGCGATAACTTGTGGCTCGACGCGCGCTCTCGGAC 304
Dbb 79 ArgValAlaAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
Qy 305 GAACACTAGTCGCGCGTGGCGATGCCACCGAGCTGGCGAGTGAGCTGTTCGACCGGCGA 364
Dbb 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
Qy 365 CGTCGCTGCTGGCGCGGAGCTGTGCACGCTGTGGATCTCTTCGACGCGGAGCCTGT 424
Dbb 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
Qy 425 CTGTGCTGCGCGCGCGCTCGGCAAGCTCGCGCCATCGCCCTCGGCGCGCGCGCGGC 484
Dbb 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155
Qy 485 ATCACGCGCACCTGCAGCACACGCTCGCACCGCGCGCGCTCGTGTCTCATGATC 544
Dbb 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
Qy 545 GCGCTCGCGCGGTGGCGCTCGCGCTCATCGCCCTCGGCGCGCTCTTGGCGGCGGC 604
Dbb 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
Qy 605 GAGGTGTGCGAGCGCTCGGCTCCAGCGCTGCCAGGTGAGCGGGGAACCTCTATGCCGCC 664
Dbb 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
Qy 665 TTCTCCACCGCGCGCGCTTCCACCTCGCGCTGGCGTGGTGGCTGTGTACCGGAAG 724
Dbb 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValValLeuPheValTyrTrpLys 235
Qy 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGCGAGAGCTGTGTCGCTCCG 784
Dbb 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValAlaValProLeuPro 255
Qy 785 GCCACCATGAGGTGAGTCCAGGTAAAGGAACACCTGATGAGGTGAAGTGGTTC 844
Dbb 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
Qy 845 ACGGCACATTGCAAGCAACAGCTGTCTTCCAGGTGAGCGGGGACTCTTGGCGGAGCAG 904
Dbb 272 ThrAlaArgArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
Qy 905 AAGGAGGCGGAGCAGCATGCTGGGAATCTTGATGCGCTGTGTGCTGCTGCTGG 964
Dbb 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
Qy 965 ATCCCTCTTCTCTGACGGAACCTCATCAGCCCACTCTGTGCTGCGAGCTGCCCCCATC 1024
Dbb 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
Qy 1025 TGGAAAAGCATATTCTGTGCTTGGCTTACTCAATTCTTCTCAACCCCTGATTAC 1084
Dbb 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
Qy 1085 ACAGCTTTTAAACAGAACTACAATGCTCTTCAAGACCTCTTACTAAGCAGAGA 1141
Dbb 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 3
SHSA_HUMAN
ID SHSA_HUMAN STANDARD; PRT; 357 AA.
AC P47898;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5-hydroxytryptamine 5A receptor (5-HT-5A) (Serotonin receptor)
DN (5-HT-5).
GN HTR5A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95080386; PubMed=7998681;
 RA Rees S., den Baas I., Foord S., Goodson S., Bull D., Kilpatrick G.,
 RA Lee M.;
 RT "Cloning and characterisation of the human 5-HT_{2A} serotonin
 RL receptor.";
 RL FEBS Lett. 355:242-246 (1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Puhl H.L. III, Ikeda S.R., Aronstam R.S.;
 RT "cDNA clones of human proteins involved in signal transduction
 RL submitted by the Guthrie cDNA resource center (www.cdna.org).";
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: X81411; CAA57168.1;
 DR EMBL: X81412; CAA57168.1; JOINED.
 DR EMBL: AF498985; AAM21132.1; -.
 DR PIR: I37107; I37107.
 DR Genbank: HGNC:5300; HTR5A.
 DR MIM: 601305; -.
 DR GO: GO:0005887; C: integral to plasma membrane; TAS.
 DR GO: GO:0004993; F: serotonin receptor activity; TAS.
 DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. .; TAS.
 DR InterPro: IPR000276; GPCR_Rhodopsin.
 DR Pfam: PF00001; 7tm 1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsin.
 DR PROSITE: PS00237; G-PROTEIN RECP F1_1; 1.
 DR PROSITE: PS00262; G-PROTEIN RECP F1_2; 1.
 DR PROSITE: PS00262; G-PROTEIN RECP F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 63 1 (POTENTIAL).
 FT DOMAIN 64 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 79 99 2 (POTENTIAL).
 FT DOMAIN 100 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 137 3 (POTENTIAL).
 FT DOMAIN 138 158 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 159 181 4 (POTENTIAL).
 FT DOMAIN 182 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 219 5 (POTENTIAL).
 FT DOMAIN 220 282 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 283 303 6 (POTENTIAL).
 FT DOMAIN 304 320 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 321 341 7 (POTENTIAL).
 FT DOMAIN 342 357 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 120 192 BY SIMILARITY.
 SQ SEQUENCE 357 AA; 40255 MW; 92F8A78C69169790 CRC64;
 Alignment Scores:
 Pred. No.: 2,19e-47 Length: 357
 Score: 1063.50 Matches: 226
 Percent Similarity: 72.98% Conservative: 36

Best Local Similarity: 62.95% Mismatches: 86
 Query Match: 49.37% Indels: 11
 DB: 1 Gaps: 6
 US-09-976-782-15 (1-1152) x SH5A_HUMAN (1-357)
 QY 65 ACCAGCAGCGAGCCCGGAGCCCAAGCCGAGAGGAGGATACCTCGTTTCGACCCCGAGCGGC 124
 Db |||||
 Db 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyLysAsp 27
 QY 125 GCCGTCCTCGCGGCGGAGCGGCCCTCTCTGCTTTCACGGTCTCGTGTGGAGCGCTG 184
 Db :::: |||||
 Db 28 AspLeuArgProSer---SerProLeuLeuSerValPheGlyValLeuLeuLeuThrLeu 46
 QY 185 CTAGTGTCTGATCGTCCGCTTCTGTCGAACTCTGTCGCTCGCTCGCTCGCTCGCTCG 244
 Db |||||
 Db 47 LeuGlyPheLeuValAlaAlaThrPheAlaThrPheAlaThrPheAlaThrPheAlaThr 66
 QY 245 CGGTCCTCGCTTCCACCGCGCTGCGGATTAATCTGTCGCTCGCTCGCTCGCTCGCTCG 304
 Db |||||
 Db 67 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86
 QY 305 GAACCTAGTGGAGCGCTGCGGATGACCGAGCTGCGGAGTGGAGTGGAGTGGAGTGGAG 364
 Db |||||
 Db 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTGCTCGCGCGGAGCTGTCACCGTGGATCTCTTCGACCGCGGAGCTGT 424
 Db |||||
 Db 106 ArgTrpGlnLeuGlyArgLeuGlyGlnLeuTrpLeuAlaCysAspVal----- 122
 QY 425 CTGTCGTCGCGCGCGCTCGGGAACGTGGCGGCGCATCGCTCGGCGCGGAGCGCGGCG 484
 Db |||||
 Db 123 LeuCysCysThrAlaSerileTrpAsnValThrAlaLeuAlaLeuAspArgTyrTrpSer 142
 QY 485 ATCAGCAGCGCCTGACGACAGCTGGCGACCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTG 544
 Db |||||
 Db 143 IleThrArgHisMetGluTyrThrLeuArgThrArgLysCysValSerAsnValMetIle 162
 QY 545 GCGCTCGCGCGCGCTCGGTCGCGCTCATGCGCGCTCGCGCGCTCGCTCGCTCGCGCGG 604
 Db |||||
 Db 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProleuLeuPheGlyTrpGly 182
 QY 605 GAGGTGTGCGAGCTCGGCTCCAGGTCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCG 664
 Db |||||
 Db 183 GluThrTyrSerGluGlySerGluGluCysGlnValSerArgGluProserTyrAlaVal 202
 QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCTGCGGTGGGTGGGTGGGTGGGTGGGTGGGTG 724
 Db |||||
 Db 203 PheSerThrValGlyAlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLys 222
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCG---CGCGCGCGAGAGCTGCTGCGCGTGG 781
 Db |||||
 Db 223 IleTyrLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242
 QY 782 CCGGCCACCATGACGTCGAGTCCCAAGGTAAAGGAGCACCTGATGAGGTGAAGTGGTG 841
 Db :::: |||||
 Db 243 SerGluAlaValGluValLysAspSerAlaLys-----GlnProGlnMetVal 258
 QY 842 TTCAGGCGACATGCAACGACGCTGCTTCGAGGTGAGCGGAGCTCTCGCGCGGAG 901
 Db |||||
 Db 259 PheThrValArg---HisAlaThrValThrPheGlnProGluGlyAspThrTrpArgGlu 277
 QY 902 CAGAAGGAGAGCGGAGCGAGCATGATGTTGGGAAATCTGATGGCGCTGTTTGTGCTGTC 961
 Db |||||
 Db 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297
 QY 962 TGGATCCCTTCTTCTCGAGGAACTCATCAGCCCACTGTGCTGCTGCTGCTGCTGCTGCTG 1021
 Db |||||
 Db 298 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317
 QY 1022 ATCTCGMAAAGCATATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081
 Db |||||
 Db 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProleuIle 337

QY 1082 TACACAGCTTTTACAGAACTACAAATGCTTCAAGACCTCTTACTAAGCAG 1138
 Db 338 TyrThrAlaPheAsnLysAsnTyrAsnSerAlaPheLysAsnPheSerArgGln 356

RESULT 4

5H5A_MOUSE
 ID 5H5A_MOUSE STANDARD; PRT; 357 AA.
 AC P30966;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE 5-hydroxytryptamine 5A receptor (5-HT-5A) (Serotonin receptor)
 DE (5-HT-5).
 GN HTR5A OR 5HT5A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RC MEDLINE=93099851; PubMed=4464308;
 RA Plassat J.-L., Boscher U., Anlaik N., Hen R.;
 RA "The mouse 5HT5 receptor reveals a remarkable heterogeneity within
 RT the 5HT1D receptor family";
 RL EMO J. 11:4779-4786(1992).
 RN [2]
 RP CHARACTERIZATION.
 RC TISSUE=Brain;
 RC MEDLINE=93196607; PubMed=9450829;
 RA Matthes H., Boscher U., Anlaik N., Grailhe R., Plassat J.-L.,
 RA Muscatelli F., Mattei M.-G., Hen R.;
 RT "Mouse 5-hydroxytryptamine 5A and 5-hydroxytryptamine 5B receptors
 RT define a new family of serotonin receptors: cloning, functional
 RT expression, and chromosomal localization.";
 RL Mol. Pharmacol. 43:313-319(1993).
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Expressed predominantly in the central nervous
 CC system; in the cerebral cortex, hippocampus, habenula, olfactory
 CC bulb and granular layer of the cerebellum.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z18278; CAA79155.1; -.
 DR MGD; MGI:96283; Htr5a.
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm1.1;
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS02662; G_PROTEIN_RECEP_F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 63 1 (POTENTIAL).
 FT DOMAIN 64 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 79 99 2 (POTENTIAL).
 FT DOMAIN 100 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 137 3 (POTENTIAL).
 FT DOMAIN 138 158 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 159 181 4 (POTENTIAL).

FT DOMAIN 182 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 219 5 (POTENTIAL).
 FT DOMAIN 220 282 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 283 303 6 (POTENTIAL).
 FT DOMAIN 304 320 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 321 341 7 (POTENTIAL).
 FT DOMAIN 342 357 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 120 192 BY SIMILARITY.
 SQ SEQUENCE 357 AA; 40804 MW; 5F5D856AC477BFAC CRC64;
 Alignment Scores:
 Pred. No.: 2,67e-45 Length: 357
 Score: 1022.50 Matches: 219
 Percent Similarity: 71.11% Conservative: 37
 Best Local Similarity: 60.83% Mismatches: 93
 Query Match: 47.47% Indels: 11
 DB: 1 Gaps: 6
 US-09-976-782-15 (1-1152) x 5H5A_MOUSE (1-357)
 QY 65 ACCAGACGGACCCCGGACCCCAAGCCGAGAGGGATATCTGGTTCCACCCCGAGCGGC 124
 Db 8 ThrSerPheSerLeuSerThrProSerSerLeuGluProAsnArgSerLeuAspThrGlu 27
 QY 125 GCCGTCTCGCGGCGGAGGCCCTTCTCTCTCTTCACGGTCTCGGTGTCGACGCTG 184
 Db 28 ValLeuArgProSerArg---ProPheLeuSerAlaPheArgValLeuValLeuThrLeu 46
 QY 185 CTAGTGTCTGATCGCTGCCACCTTCTCTGTGGAACTCTGCTGGTTCCTCCGTCACCATCCG 244
 Db 47 LeuGlyPheLeuAlaAlaAlaThrPheThrTrpAsnLeuValLeuAlaThrLeuLeu 66
 QY 245 CGGTCTGCTGCTTCCACCGCGTCCCGGATAACTGTGTGGCTCGACGGCGCTCTCGAC 304
 Db 67 LysValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaIleSerAsp 86
 QY 305 GAATAGTGGAGCGCTGGCGATGCCACCGACCTGGCGAGTGGAGTGTCTGACCGGGCGA 364
 Db 87 ValLeuValAlaValLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTGCTGGCGGAGCTGTGCCACCTGTGGATCTCTCTCGACGCCGAGCGCTGT 424
 Db 106 ArgTrpGlnLeuGlyArgArgLeuGlyCysGlnLeuTrpIleAlaCysAspVal----- 122
 QY 425 CTGTCTGCCCGCGCGCTCGGAAACGTGGCGGCATCGCCCTGGCGCGGCGGCGGCGCC 484
 Db 123 LeuCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSer 142
 QY 485 ATCACGGGACCTGCAGCACACCTGCGACCCGCGCGCGGCTCGTTCATGATGTC 544
 Db 143 IleThrArgHisLeuGluTyrThrLeuArgThrArgLysArgValSerAsnValMetIle 162
 QY 545 CGCTGCGCGCGGCTCGCGCTCATCCCTCGCGCGCTGCTCTTTGGCGCGGCGC 604
 Db 163 LeuLeuThrTrpAlaLeuSerThrValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 QY 605 GAGTGTGCGAGCTCGGCTCCAGCTGCCAGTGCAGGTGAGCGCGGGAACCTCTATGCCGCC 664
 Db 183 GluThrTyrSerGluProSerGluCysGlnValSerArgGluProSerTyrThrVal 202
 QY 665 TTCTCCACCGCGGCGCTTCCACTGCGGCTTGGCGTGGTGGCGTGTGTGTACCGGAG 724
 Db 203 PheSerThrValGlyAlaPheTyrLeuProLeuTrpLeuValLeuPheValTyrTrpLys 222
 QY 725 ATCTACGAGGCGCAAGTTCTTTCGGC---CGCGCGGAGAGCTGTGCTCGCGTGTG 781
 Db 223 IleTyrArgAlaAlaLysPheArgMetGlySerArgLysThrAsnSerValSerProVal 242
 QY 782 CCGGCCACCATGAGGTGAGGTCCCAAGGTAAAGGACCACTGATGAGGCGTGAAGTGTGTG 841
 Db 243 ProGluAlaValGlu-----ValLysAsnAlaThrGlnHisProGlnMetVal 258

QY 842 TTCAAGGCACATTGCAAGCAAGCGTGTCTCTCCAGGTGAGCGGAGCTCTGCGCGGAG 901
 DB 259 PheThrAlaArg--HisAlaThrValThrPheGlnThrGluGlyAspThrTrpArgGlu 277
 QY 902 CAGAAGGAGGAGGAGCAGCAGCATGATGGTGGGAATCTGATGGCGTGTGTGGTGTGC 961
 DB 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyLeuIleGlyValPheValLeuGly 297
 QY 962 TGGATCCCTCTCTCTGAGGGAACATCATCAGCCCACTCTGTGCTGCGAGCCTGCCCC 1021
 DB 298 TrpPheProPheValThrGluLeuIleSerProLeuGlySerTrpAspValProAla 317
 QY 1022 ATCTGGAAGAGCATATTTCTGTGGCTTGGCTACTCCAAATCTTTCTTCAACCCCTGAT 1081
 DB 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIle 337
 QY 1082 TACACAGCTTAAACAGAACTACAAACATGCTTCAAGAGCCTCTTTACTAAGCAGAGA 1141
 DB 338 TyrThrAlaPheAsnArgSerTyrSerSerAlaPheLysValPhePheSerLysGlnGln 357
 RESULT 5
 ID SH5A_RAT STANDARD; PRT; 357 AA.
 AC P35364;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)
 DE 5-hydroxytryptamine 5A receptor (5-HT-5A) (Serotonin receptor)
 DE (RC17).
 GN HTR5A OR 5HT5A.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=93234515; PubMed=7682702;
 RA Erlanger M.G., Lovenberg T.W., Baron B.M., de Lecea L.,
 RA Danielson P.E., Racke M., Slone A.L., Siegel B.W., Foye P.E.,
 RA Cannon K., Burns J.E., Sutcliffe G.J.,
 RT "two members of a distinct subfamily of 5-hydroxytryptamine receptors
 RT differentially expressed in rat brain."
 RL Proc. Natl. Acad. Sci. U.S.A. 90:3452-3456 (1993).
 CC -!- FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: Central nervous system.
 CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC STRONGEST TO THE OTHER 5HT-5 SUBTYPE RECEPTORS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L10072; AAA40615.1; -
 CC PIR; B47472; B47472.
 DR InterPro; IPR00276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm 1; 1.
 DR PRINTS; PR00237; GPCR_RHODPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR F1_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family.
 FT DOMAIN 1 40 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 41 63 1 (POTENTIAL).
 FT

FT DOMAIN 64 78 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 79 99 2 (POTENTIAL).
 FT DOMAIN 100 115 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 116 137 3 (POTENTIAL).
 FT DOMAIN 138 158 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 159 181 4 (POTENTIAL).
 FT DOMAIN 182 198 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 199 219 5 (POTENTIAL).
 FT DOMAIN 220 282 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 283 303 6 (POTENTIAL).
 FT DOMAIN 304 320 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 321 341 7 (POTENTIAL).
 FT DOMAIN 342 357 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 120 192 BY SIMILARITY.
 SQ SEQUENCE 357 AA; 40672 MW; 8C498A50C88408B5 CRC64;
 Alignment Scores:
 Pred. No.: 6.8e-45 Length: 357
 Score: 1014.50 Matches: 216
 Percent Similarity: 71.27% Conservative: 37
 Best Local Similarity: 60.85% Mismatches: 81
 Query Match: 47.10% Indels: 21
 DB: 1 Gaps: 7
 US-09-976-782-15 (1-1152) x SH5A_RAT (1-357)
 QY 110 TCGACCCGAGCGGCCCTCTCTGCGCGCGGAGCGGCCGCC----- 151
 DB 13 SerThrProSer---ThrLeuGluProAsnArgSerLeuAspThrGluAlaLeuArgThr 31
 QY 152 -----TTCTCTGCTTTCACGTCCTGCTGGTGGTGCAGCTGCTAGTGTGCTGATC 199
 DB 32 SerGlnSerPheLeuSerAlaPheArgValLeuValLeuLeuLeuGlyPheLeuAla 51
 QY 200 GCTGCCACTTCTGTGGAACTGCTGTTCCGGTCCACATCCCGGGTCCGTGCTTTC 259
 DB 52 AlaAlaThrPheThrTrpAsnLeuValLeuAlaThrIleLeuArgValArgThrPhe 71
 QY 260 CACCGGTGCGCATAACTTGTGGCTGCGGCGCTCGGACGAACCTAGTGGCAGCG 319
 DB 72 HisArgValProHisAsnLeuValAlaSerMetAlaIleSerAspValLeuValAlaVal 91
 QY 320 CTGGCGAGTCCACCGAGCTGCGAGTGTGACCGGCGAGCTCGGCTGCTGGCG 379
 DB 92 LeuValMetProLeuSerLeuValHisGluLeuSer---GlyArgArgTrpGlnLeuGly 110
 QY 380 CGAGCTGTGCGACGTGTGGATCTCTTCGACGCGGAGCCTGTCTGTCTGCCCGCC 439
 DB 111 ArgArgLeuCysGlnLeuTrpIleAlaCysAspVal-----LeuCysCysThrAla 127
 QY 440 GGCCTCGGAAGCTGGCGCCATCCCTGCGCGCGGAGCGGCCCATCACAGGCACTG 499
 DB 128 SerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSerIleThrArgHisLeu 147
 QY 500 CAGCACAGCTCGCACCGCAGCGCGCTGCTGTGCTCATCATCGCTCGCCCGGGG 559
 DB 148 GluTyrThrLeuArgAlaArgLysArgValSerAsnValMetIleLeuLeuThrTrpAla 167
 QY 560 CGCTCGGCGCTCATCGCCCTCGCGCGCTGCTTTTGGCGCGGAGGTGTGCGACGCT 619
 DB 168 LeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGlyGluThrTyrSerGlu 187
 QY 620 CGGCTCCAGCGCTGCAGTGCAGCGGAGCAACCTCTATGCGCCTCTCCACCGCGCG 679
 DB 188 LeuSerGluGluCysGlnValSerArgGluProSerTyrThrValPheSerThrValGly 207
 QY 680 GCCTTCCACCTGCGCGCTGGCGTGTGCTGTCTACCGGAAGATCTTACGAGCGCGC 739
 DB 208 AlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLysIleTyrLysAlaAla 227
 QY 740 AAGTTTCTGTTTCGGC---CGCGCGGAGAGCTGTGCTGCCGCTTGCCTGCCACCATGCAG 796

DR EMBL; M89478; -, NOT ANNOTATED_CDS.
 DR EMBL; L09732; AAA36030.1; -
 DR EMBL; M83180; AAA36029.1; -
 DR EMBL; AB041370; BAA94455.1; -
 DR EMBL; AY252227; AAO67712.1; -
 DR EMBL; AL049595; CAB51537.1; -
 DR PIR; JN0268; JN0268.
 DR Genew; HGNC:5287; HTRIB.
 DR MIM; 182131; -
 DR GO; GO:0005887; C-integral to plasma membrane; TAS.
 DR GO; GO:004993; F-serotonin receptor activity; TAS.
 DR GO; GO:0007187; P-G-protein signaling, coupled to cyclic nucl. .; TAS.
 DR GO; GO:0007268; P-synaptic transmission; TAS.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm.1; 1
 DR PRINTS; PR00237; GPCRHOOPS.
 DR PROSITE; PS00237; G PROTEIN RECP F1_1; 1.
 DR PROSITE; PS0262; G PROTEIN RECP F1_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate; Phosphorylation;
 KW Polymorphism.
 FT DOMAIN 1 49 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 50 73 1 (POTENTIAL).
 FT DOMAIN 74 86 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 87 109 2 (POTENTIAL).
 FT DOMAIN 110 119 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 120 145 3 (POTENTIAL).
 FT DOMAIN 146 165 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 166 187 4 (POTENTIAL).
 FT DOMAIN 188 205 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 206 229 5 (POTENTIAL).
 FT DOMAIN 230 315 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 316 340 6 (POTENTIAL).
 FT DOMAIN 341 347 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 348 373 7 (POTENTIAL).
 FT DOMAIN 374 390 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 24 24 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 32 32 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT DISULFID 122 199 BY SIMILARITY.
 FT LIPID 388 388 S-palmitoyl cysteine (Potential).
 FT VARIANT 124 124 F -> C (in dbSNP:130060).
 FT VARIANT 219 219 F -> L (in dbSNP:130061).
 FT VARIANT 367 367 I -> V (in dbSNP:130063).
 FT VARIANT 374 374 E -> K (in dbSNP:130064).
 FT SEQUENCE 390 AA; 43568 MW; CD874DC7EB44CF12 CRC64;
 Alignment Scores:
 Pred. No.: 4 31e-16 Length: 390
 Score: 448.00 Matches: 127
 Percent Similarity: 48.35% Conservative: 63
 Best Local Similarity: 32.32% Mismatches: 147
 Query Match: 20.80% Indels: 56
 DB: 1 Gaps: 13
 US-09-976-782-15 (1-1152) x SH1B_HUMAN (1-390)
 QY 59 CCGAGACCGAGCGGACCCGAGCCGAGAGGGATACGCGTTGACCCCG 118
 Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnSerSerAlaPro 29
 QY 119 AGC-----GGCGCCCTCTCGCGGCCGA 142
 Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
 QY 143 GGGCGCCCTCTCTCTGCTGCTACGGTCCTGGTGACGCTGCTAGTGTGCTGCTGCT 202
 Db 48 -----TrpLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
 QY 203 GCACCTTCTGTGGAACTGCTGTTCCGGTCCACCATCCCGGGGTCGTCCTCCAC 262

Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
 QY 263 CGCGTCCGCATAACTTGTGTCGACGGCGCTCGAGCAACTAGTGGCAGCGCTG 322
 Db 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
 QY 323 CGGATGCCACCGAGCTGGCGAGTGTGCGACGGCGGCGAGCTGGCTGGTGGCGCG 382
 Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
 QY 383 AGCTGTGCGACGTGTGATCTCTTCGACGGCGGAGCCGTGTGCTGCTGCTGCCCGCGGC 442
 Db 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
 QY 443 CTCGGGAACGTGGCGCCATCGCCCTGGCGCGGCGGCGGCGCATCACAGGCACTGCGAC 502
 Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
 QY 503 CACAGCTGCGCACCGCGAGCGCGCTGCTGCTCATGATCGCGCTCGCGCGGTGCGC 562
 Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
 QY 563 TCGGCGCTCATCCCTCGCGCGCTCTTTTGGCGGGCGAGGTGTCGACGCTCGG 622
 Db 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGlu 195
 QY 623 CTCGACGCTGCCAGTGTGAGCGGGAACCC---TCCTATGCGCGCTTCTCCACCGCGGC 679
 Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
 QY 680 GCCTTCCACCTGCGCGTGGCGTGTGCTACCGGAAGATCTACGAGGCGGC 739
 Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
 QY 740 AAGTTTCGTTTC-----GGCGCGGC---CGGAGAGCTGTG 772
 Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
 QY 773 CTG-----CCGTGCGCGCCACCATCGAGTGTGAGGTCCAGGTAAAG 814
 Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
 QY 815 GAAGCACCTGATGAGCT---GAAGTGTGTGTTCACGGCACATTGCAAGCAACGCTGTC 871
 Db 276 AspValProSerGluSerGlySerProValTyrValAsnGlnValLysValArgValSer 295
 QY 872 TTCCAGGTGAGCGGACTCTCTGGCGGAGCAGAG-----GAGAGG 913
 Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
 QY 914 CGAGCAGCATGATGTTGGGAATTCGATTGGCGTGTGTGCTGTGCTGATCCCTTC 973
 Db 311 LysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
 QY 974 TTCCTGACGAACTCATCAGCCACTCTGT-----GCCTGACGCTGCGCCCATCTGG 1027
 Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
 QY 1028 AAAAGCATTTCTGTGCTTGGCTTACTCCAAATCTTTCTTCAACCCCTGATTTACACA 1087
 Db 351 PheAppPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThr 370
 QY 1088 GCTTTTAAACAAGAACTACAAACATGCTTCAAGAGCCTC 1126
 Db 371 MetSerAsnGluAspPheLysGlnAlaPheHisLysLeu 383
 RESULT 7
 SH1B_PANTR
 ID SH1B_PANTR STANDARD; PRT; 390 AA.
 AC P6020;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)

RESULT 8
 5H1B_SPAEH STANDARD; PRT; 386 AA.
 ID 5H1B_SPAEH
 AC P56496;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor).
 GN HTR1B OR 5H1B.
 OS Spalax leucodon ehrenbergi (Ehrenberg's mole rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Spalacinae;
 OC Nanospalax.
 OX NCBI_TaxID=30637;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Devor E.J., Nevo E.;
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
 CC FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins that inhibit adenylyate
 CC cyclase activity.
 CC SUBCELLULAR LOCATION: Integral membrane protein.
 CC SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Strongest to the other 5HT-1 subtype receptors.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF027184; AB82748.1; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm 1; 1.
 CC PRINTS; PR00237; GPCR_Rhodopsn.
 CC PROSITE; PS00237; G-PROTEIN RECEPTOR FL_1; 1.
 CC PROSITE; PS0262; G-PROTEIN RECEPTOR FL_2; 1.
 CC G-protein coupled receptor; Transmembrane; Glycoprotein;
 CC Multigene family; Lipoprotein; Palmitate.
 KW DOMAIN 1 47 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 48 71 1 (POTENTIAL).
 FT DOMAIN 72 82 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 83 109 2 (POTENTIAL).
 FT DOMAIN 110 119 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 120 141 3 (POTENTIAL).
 FT DOMAIN 142 163 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 164 183 4 (POTENTIAL).
 FT DOMAIN 184 204 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 205 224 5 (POTENTIAL).
 FT DOMAIN 225 311 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 312 332 6 (POTENTIAL).
 FT DOMAIN 333 346 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 347 369 7 (POTENTIAL).
 FT DOMAIN 370 386 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 24 24 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT CARBOHYD 28 28 N-LINKED (GLNAC. . .) (POTENTIAL).
 FT DISULFID 118 195 BY SIMILARITY.
 FT LIPID 384 394 S-palmitoyl cysteine (POTENTIAL).
 SQ SEQUENCE 386 AA; 43101 MW; 7A006E021A4B7F4 CRC64;

Alignment Scores:

Pred. No.: 4,84e-16 Length: 386
 Score: 447.00 Matches: 129
 Percent Similarity: 47.85% Conservative: 60
 Best Local Similarity: 32.66% Mismatches: 142
 Query Match: 20.75% Indels: 64
 DB: 1 Gaps: 14

US-09-976-782-15 (1-1152) x 5H1B_SPAEH (1-386)
 QY 77 CCGGGACC-----CCAAAGCCGAGAGGATACCTCGTTTCGACCCGAGC--- 121
 Db 4 ProGlyAlaArgCysAlaProProProAlaGlySerGlnThrGlnThrProSerSer 23
 QY 122 -----GGCCCGCTCTCTGCGC 136
 Db 24 AsnLeuSerHisAsnCysSerAlaAspSerTyrlleTyrlleGlnAspSerileAlaLeuPro 43
 QY 137 GCGGAGGGCCGCCCTTCTCTCTTTCAGGTCCTGGTGTGAGCGTGTAGTGTCTGCTG 196
 Db 44 -----TrpLysValLeuLeuValAlaLeuAlaLeuAlaLeu 55
 QY 197 ATCGCTGCCACTTCTCTGTGGAACTGCTGCTCCGTCCACCATCCCGGGTCCCTGCC 256
 Db 56 ThrLeuAlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrlleThrArgThrArgLys 75
 QY 257 TTCCACCGCGTCCGATTAACCTTGTGGCTCGACGGCGCTCTCGAGCAAGTCTAGTGGCA 316
 Db 76 LeuHisThrProAlaAsnTyrlleAlaSerLeuAlaValThrAspLeuLeuValSer 95
 QY 317 GCGCTGGCATGCCACCGAGCTGCGAGTGTGTGCGCGGCGAGTGTGGTGTCTGCTG 376
 Db 96 IleLeuValMetPro-----IleSerMetTyrlleThrValThrGlyArgTrpThrLeu 113
 QY 377 GCGCGAGCTGTGTGACGTGTGATCTCTTCGACGCGGAGCTGTCTGTCTGCCCC 436
 Db 114 GlyGlnValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThr 130
 QY 437 GCGCGCTCGGAAGCTGTGGCCCATCGCCCTGGCGCGGCGGCGGCGGCGGCGGCGGCA 496
 Db 131 AlaSerIleMetHisLeuCysValIleAlaLeuAspArgTyrlleAlaIleThrAspAla 150
 QY 497 CTGAGCACACCTGCGGACCGCGGCGGCTGTGCTGTATGATCGGCTGTGCGCGG 556
 Db 151 ValGluTyrlleSerAlaLysArgThrProArgArgAlaAlaValMetIleAlaLeuValTrp 170
 QY 557 GTGCGGTGCGGCTCATCGCTCGCGCGCTCTTTGGCGGCGGCGGCGGCGGCGGCGG 616
 Db 171 ValPheSerIleSerIleSerLeu---ProArgPhePheTrpArgGlnAlaLysAlaGlu 189
 QY 617 GCTCGGCTCACGCTGTGAGTGTGAGCGCGGAA---CCCTCTATGCGCGCTTCTCCAC 673
 Db 190 GluGluValLeuAspCysLeuValAsnThrAspHisValLeuTyrlleThrValTyrlle 209
 QY 674 CCGCGGCTTCCACTGCGGCTGTGCGGTGCGGCTGTGCTACCGAGATCTACGAG 733
 Db 210 ValGlyAlaPheTyrlleProThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuVal 229
 QY 734 GCGGCCAAGTTTCGTTC-----GGCGCGCGC---CGAGA 766
 Db 230 GluAlaArgSerArgIleLeuLysGlnThrProAsnLysThrGlyLysArgLeuSerArg 249
 QY 767 GGTGTGCTG-----CCGTTGCGGCGCACCATGCGAGTGTGAGTGTGAGTGTGAG 808
 Db 250 AlaGlnLeuIleSerAspSerProGlySerThrSerSerValThrSerIleAsnSerArg 269
 QY 809 GTAAGGAGACCTCATGAGGCT---GAAGTGTGTTCACCGGCACATTGCAAGCAACG 865
 Db 270 ValProAspValProSerGluSerGlySerProValTyrlleValAsnGlnValLysValArg 289
 QY 866 GTGTCTTCTTCAGTGTGCGGAGCTCTGCGGCGGAGCAAG----- 907
 Db 290 ValSer-----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArg 304
 QY 908 GAGAGCGAGCGCATGATGTGGAGTCTGATGGGTGTGCTGTGCTGTGCTGTGCTGTG 967
 Db 305 GluArgLysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleValCysTrpLeu 324
 QY 968 CCTTCTTCTGCGGAACTCATCAGCCCACTGT-----GCTGTGAGCGCTGTGCGGCGG 1021
 Db 325 ProphePheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisMet 344

QY 1022 ATCTGGAAGACATATTCTGTGGCTTGCTACTCCAAATCTTTCTTCAACCCCTGATT 1081
 Db 345 AlallepheasphepneantrpLeuGlyTyrLeuAsnSerLeuLeuAlaAsnProIle 364

QY 1082 TACACAGCTTTTAAACAGAACTACAAATCGCTTCAAGAGCCTC 1126
 Db 365 TyrThrMetProAsnGluaspPheLysGlnAlaPheHisLysLeu 379

RESULT 9
 SHIB_CAVPO STANDARD; PRT; 389 AA.
 AC O08832;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE 5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor)
 DE (5-HT1B receptor).
 GN HTR1B.
 OS Cavia porcellus (Guinea pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
 OX NCBI_TaxID=10141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=97368662; PubMed=9225276;
 RA Zgombick J.M., Bard J.A., Kucharewicz S.A., Urquhart D.A.,
 RA Weinschank R.L., Branchek T.A.;
 RT "Molecular cloning and pharmacological characterization of guinea pig
 RT 5-HT1B and 5-HT1D receptors.";
 RL Neuropharmacology 36:513-524 (1997).
 CC -1- FUNCTION: This is one of the several different receptors for 5-
 CC hydroxytryptamine (serotonin), a biogenic hormone that functions
 CC as a neurotransmitter, a hormone, and a mitogen. The activity of
 CC this receptor is mediated by G proteins that inhibit adenylyate
 CC cyclase activity.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
 CC Strongest to the other 5HT-1 subtype receptors.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U82175; AAB58500.1; -;
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00061; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G PROTEIN RECF F1 1; 1.
 DR PROSITE; PS0262; G PROTEIN RECF F1 2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Multigene family; Lipoprotein; Palmitate.
 KW
 FT DOMAIN 1 50 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 51 74 1 (POTENTIAL).
 FT DOMAIN 75 85 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 86 112 2 (POTENTIAL).
 FT DOMAIN 113 122 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 123 144 3 (POTENTIAL).
 FT DOMAIN 145 166 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 167 186 4 (POTENTIAL).
 FT DOMAIN 187 207 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 208 227 5 (POTENTIAL).
 FT DOMAIN 228 314 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 315 335 6 (POTENTIAL).
 FT DOMAIN 336 349 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 350 372 7 (POTENTIAL).
 FT DOMAIN 373 389 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 24 24 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT

FT CARBOHYD 31 31 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 121 198 BY SIMILARITY.
 FT LIPID 387 387 S-palmitoyl cysteine (Potential).
 SQ SEQUENCE 389 AA; 43110 MW; C057CAB0A7FEE3C6 CRC64;

Alignment Scores:
 Pred. No.: 6,49e-16 Length: 389
 Score: 444.50 Matches: 125
 Percent Similarity: 48.44% Conservative: 61
 Best Local Similarity: 32.55% Mismatches: 159
 Query Match: 20.64% Indels: 39
 DB: 1 Gaps: 12

US-09-976-782-15 (1-1152) x SHIB_CAVPO (1-389)

QY 59 CCGAGACAGCAGCGGACCGGAGCCCAAGCCCGAGAGGATCTCGGTTCGACCCCG 118
 Db 10 ProProAlaValLeuGlySerGlnThrGlyLeuProHisAlaAsnValSerAlaProPro 29

QY 119 AGCGGCGCGCTCTGCGGCGCGGAGGCGCGCTTCTCTGTC-----TTCACGGTC 169
 Db 30 AsnAsnCysSerAlaProSerHisIleTyrGlnAspSerIleAlaLeuProTyrLysVal 49

QY 170 CTGTGTGTGACGCTGTAGTCTGCTGATCGTCCACTTCTCTGTGTGGAACCTGCTGTT 229
 Db 50 LeuLeuValValLeuLeuAlaLeuLeuThrLeuAlaThrThrLeuSerAsnAlaPheVal 69

QY 230 CCGGTACCATCCGCGGGTCCGTGCTTCCACCGCGTCCGCGGATACTTGGTGGCCCTCG 289
 Db 70 IleAlaThrValTyrArgThrArgLysLeuHisThrProAlaAsnTyrLeuIleAlaSer 89

QY 290 ACGGCGCTCTCGACGAACCTAGTGGCAGCTGCGGATGCCACGACCTGGCGAGTGAG 349
 Db 90 LeuAlaPheThrAspLeuLeuValSerIleLeuValMetPro-----IleSerThrMet 107

QY 350 CTGTGCGACCGGCGACCTGCTGCGCGCGGAGCCTGTGCCACGTGTGGATCTCTCTTC 409
 Db 108 TyrThrValThrGlyArgTyrThrLeuGlyGlnAlaLeuCysAspPheTyrLeuSerSer 127

QY 410 GAGCGCGGAGCTGTGTGTCTGCTGCGCGCGCTCGGACGTGGCGGCATCGCCCTG 469
 Db 128 Asp-----IleThrCysThrAlaSerIleMethHisLeuCysValIleAlaLeu 144

QY 470 GCGCGACGCGGCGCATCACAGGCACCTGCAGCACAGCTGCGCACCCGCGAGCGCGCC 529
 Db 145 AspArgTyrTrpAlaIleThrAspAlaValGlyTyrSerAlaLysArgThrProArgArg 164

QY 530 TCGTGTCTCATGATCGCGCTCGCCCGGGTCCGCTCGCGCTCATCGCCCTCGCGCGCTG 589
 Db 165 AlaAlaGlyMetIleAlaLeuValTyrPheSerIleCysIleSerLeuProPro--- 183

QY 590 CTCTTTGGCGGCGAGGTGTGGCAGCTCGGCTCCAGCGCTCCAGAGTGAGCGGGAA 649
 Db 184 PhePheTyrArgGlnAlaLysAlaGluGluValLeuLeuAspCysLeuValAsnThrAsp 203

QY 650 ---CCCTCTATGCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCGCTTGGGTGGTGTG 706
 Db 204 HisValLeuTyrThrValTyrSerThrGlyGlyAlaPheTyrLeuProThrLeuLeuLeu 223

QY 707 CCGTTTGTCTACCGGAAGATCTACGAGCGCGCGCAAGTTTCGTTTC----- 751
 Db 224 IleAlaLeuTyrGlyArgIleTyrValGluAlaArgSerArgIleLeuLysGlnThrPro 243

QY 752 -----GGCGCGCGC---CGAGAGCTGTGCTG-----CCGTTG 781
 Db 244 AsnLysThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySerThr 263

QY 782 CCGGCCACCATGCGAGGTGAGTCCCAAGGTAAAGGAAGCACCT---GATGAGGCTGAAGT 838
 Db 264 SerSerValThrSerIleAsnSerArgAlaProGluValProCysAspSerGlySerPro 283

QY 839 GTGTTCACGGCACATTCGAAAGCAACCGTGTCTTCCAGGTGAGCGGGGACTCTCTGCGG 898
 Db 839 GTGTTCACGGCACATTCGAAAGCAACCGTGTCTTCCAGGTGAGCGGGGACTCTCTGCGG 898


```

Qy 467 CTGGCGCGCGAGCGGCATCACAGCGCACTGCAGCACACGCTGGCGCACCCGCGAGCGCG 526
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
145 LeuAspArgTyrTrpAlaIleThrAspAlaValGluTyrSerAlaIysArgThrProLys 164
Qy 527 GCCTCGTGCTCARGATCGCGCTCGCCGGGTGGCGCTCATGCCCTCGCGCGCG 586
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
165 ArgAlaIleMetIleArgLeuValTrpValPheSerIleCysIleSerLeuProPro 184
Qy 587 CTGCTCTTTGGCGGGGAGGTGTGGACGCTGGCTCCAGCGCTGCACGTTGAGCGCG 646
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
185 ---PhePheTrpArgGlnAlaLysAlaGluGluValSerGluCysLeuValAsnThr 203
Qy 647 GAA---CCCTCTATGCGCGCTTCTCCACCGCGCGGCTTCCACCTGCCGCTGGCGTG 703
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
204 AspHisValLeuTyrThrValTyrSerThrValGlyAlaPheTyrLeuProThrLeu 223
Qy 704 GTGCGGTTGTCTACGGAAGATCTACGAGCGCGCAAGTTTCGTTC----- 751
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
224 LeuIleAlaLeuTyrGlyArgIleTyrValGluAlaArgSerArgIleLeuLysGlnThr 243
Qy 752 -----GGCGCGCGC---CGGAGAGCTGTGCTG-----CCG 778
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
244 ProAsnArgThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySer 263
Qy 779 TTGCGGCGCACCATGCAGTGCAGTCCCAAGGTAAAGGAGCACCTGATGAGCT---GAA 835
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
264 ThrThrSerValThrSerIleAsnSerArgAlaProAspValProSerGluSerGlySer 283
Qy 836 GTGTGTCTACGCGACATFTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACCTCGG 895
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
284 ProValTyrValAsnGlnValLysValArgValSer-----AspAlaLeu 298
Qy 896 CGGAGCGACAGAG-----CAGAGCGCGAGCGCATGATGCTGGGATTT 937
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
299 LeuGluLysLysLysLeuMetAlaAlaArgLysAlaThrLysLeuGlyIle 318
Qy 938 CTGATTGGCGGTGTGTCTGTGTCGATCCCTTCTTCGTACGAGCACTCATGAGCCCA 997
Db :::: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
319 IleLeuGlyValPheIleValCysTrpLeuProPhePheIleIleSerLeuValMetPro 338
Qy 998 CTCTGT-----GCTCGAGCGCTGCCCCCATCTGGAAGACATATTTCTGTGGCTGGC 1051
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
339 IleCysLysAspAlaCysTrpPheHisGlnAlaIlePheAspPheThrTrpLeuGly 358
Qy 1052 TACTCCATTTCTTCTCAACCCCTGATTTACAGAGCTTTTACAGAACTACACAACT 1111
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
359 TyrValAsnSerLeuIleAsnProIleIleTyrThrMetSerAsnGluAspPheLysGln 378
Qy 1112 GCCTTCAAGAGCGCTC 1126
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
379 AlaPheHisLysLeu 383

RESULT 11
SHT1 DROME
ID -SHT1 DROME STANDARD; PRT; 564 AA.
AC P20905; OSV21;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE 5-hydroxytryptamine receptor 1 (5-HT receptor) (Serotonin receptor).
GN 5-HT7 OR 5HT-R1 OR CG12073.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Oregon-R; TISSUE=Head;
RX MEDLINE=91062395; PubMed=2174167;
RA Witz P., Anlaik N., Plassat J.-L., Maroteaux L., Borrelli E., Hen R.;
RT "Cloning and characterization of a Drosophila serotonin receptor that
activates adenylylate cyclase.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 87:8940-8944 (1990).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=Berkley;
RC MEDLINE=20196006; PubMed=10731132;
RX Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
BRANDON R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
ABRILL J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
BAILEY R.M., Bagnay A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
BEESON K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
BORKOVA D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
BURTS K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
CHERRY J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
DE PABLOS B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
DODSON K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
DURBIN K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
FOSTER C., Gabriellian A.B., Garg N.S., Gelbart W.M., Glasser K.,
GLODEK A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
HARRIS N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
HOSTIN D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
JALALI M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
KIMMEL B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
LAO P., Lei Y., Levisky A.A., Li J.H., Li Z., Liang Y., Lin X.,
LIU X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
MERKULOV G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
MOUNT S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
NELSON D.M., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
PALAZZOLO M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
REINERT K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
SHUE B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
SPIER E., Spradling A.C., Stapleton M., Strong R., Sun E.,
SVIRSKAS R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
WANG Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
WILLIAMS S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
YE J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
GIBBS R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195 (2000).
CC -!- FUNCTION: This is one of the several different receptors for 5-
CC hydroxytryptamine (serotonin), a biogenic hormone that functions
CC as a neurotransmitter, a hormone, and a mitogen. The activity of
CC this receptor is mediated by G proteins which activate adenylylate
CC cyclase.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: Head.
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M55533; AAA28305.1; -.
CC EMBL; AE003776; AAF57104.1; -.
CC FlyBase; FBgn0004573; 5-HT7
CC InterPro; IPR000276; GPCR_Rhodopsn.
CC InterPro; IPR007455; Serglycin.
CC Pfam; PF00001; 7tm1; 1.
CC Pfam; PF04360; Serglycin; 1.
CC PRINTS; PR00237; GPCR_RHODOPSN
CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Glycoprotein; Repeat.
FT TRANSMEM 29 51 0 (POTENTIAL).
FT TRANSMEM 165 188 1 (POTENTIAL).

```


	836	G T G C G T C T T A C G C C A C A T T G C A A A G C A A C G G T G C T C C T T C C A G G T G A C G C G G G A C T C C T C G G	895
QY		: :	
Dd	280	P r o V a l t y r V a I a s n G l n V a l l y S v A l g V a l s e r -----A s p A l a L e u	294
		: :	
QY	896	C G G G A C A C A A G -----G A G A G C G C A G C C C A T G A T G T G T G G G A A T T	937
Ddb	295	L e u G l u L y S l y S l y S l e u M e t A l a A r g G l u a r g l y S l a T h r l y S T h r L e u G l y l e	314
		: :	
QY	938	C T G A T T G G C G T G T T G T G T G T G T G T G A T C C C C T T C T T C T G A C G G A A C A T C A T C A G C C C A	997
Ddb	315	I l e L e u G l y A l a P h e i l e V a l C y s T r p L e u P r o P h e H e i l e l l e S e r L e u V a l M e t P r o	334
QY	998	C N C T G F -----G C C T G A C G C C T G C C C C C A T C T G G A A A A G C A T A T T C T G T G G C T T G G C	1051
Ddb	335	I l e C y L y S A s p A l a c Y s T r p P h e H i s M e t A l a T h r L e u A s p P h e P h e A s n T r p L e u G l y	354
QY	1052	T A C T C C A A T T C T T T C T T C A C C C C C T G A T T T A C A G C T T T T A A C A G A A C T A C A A C A T	1111
Ddb	355	T y r L e u A s n S e r L e u l e A s n P r o l l e i l e y T r T h r M e t S e r A s n G l u A s p P h e L y s G l n	374
QY	1112	G C T T T C A A G A G C C T C	1126
Dd	375	A l a P h e H i S l y S l e u	379
RESULT 14			
SUBSEQUENCE MAP			
ID	SH1B MOUSE	STANDARD;	PRT; 386 AA.
AC	P28374;		
DT	01-DEC-1992 (Rel. 24, Created)		
DT	01-DEC-1992 (Rel. 24, Last sequence update)		
DE	10-OCT-2003 (Rel. 42, Last annotation update)		
DE	5-hydroxytryptamine 1B receptor (5-HT-1B) (Serotonin receptor).		
GN	HTR1B OR 5HT1B.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=92212959; PubMed=1557407;		
RA	Maroteaux L., Saudou F., Amlaiky N., Boschert U., Plassat J.-L.,		
RA	Ren R. ;		
RT	"Mouse 5HT1B serotonin receptor: cloning, functional expression, and		
RT	localization in motor control centers.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 89:3020-3024(1992).		
CC	-!- FUNCTION: This is one of the several different receptors for 5-		
CC	hydroxytryptamine (serotonin), a biogenic hormone that functions		
CC	as a neurotransmitter, a hormone, and a mitogen. The activity of		
CC	this receptor is mediated by G proteins that inhibit adenylylate		
CC	cyclase activity.		
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.		
CC	-!- TISSUE SPECIFICITY: Predominantly expressed in striatum and		
CC	Purkinje cells.		
CC	-!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.		
CC	Strongest to the other 5HT-1 subtype receptors.		

This SWISS-PROT entry is copyright. It is produced through a collaboration			
between the Swiss Institute of Bioinformatics and the EMBL outstation -			
the European Bioinformatics Institute. There are no restrictions on its			
use by non-profit institutions as long as its content is in no way			
modified and this statement is not removed. Usage by and for commercial			
entities requires a license agreement (See http://www.isb-sib.ch/announce/			
or send an email to licensese@isb-sib.ch).			

DR	EMBL; Z11597; CAA77678.1; --		
DR	EMBL; M85151; AAA83221.1; --		
DR	PIR; A42688; A42688.		
DR	MGD; MG1.96274; Htrlb.		
DR	InterPro; IPR000276; GPCR_Rhodpsn.		
DR	Pfam; PF00001; 7tm_1; 1.		
DR	PRINTS; PR00237; GCRRRHODPSN.		
DR	PROSITE; PS00237; G PROTEIN RECF Fl 1; 1.		
DR	PROSITE; PSS0262; G PROTEIN RECF Fl 2; 1.		

Db	78	GlySerLeuAlaThrThrAspLeuLeuValSerIleLeuValMetProIleSerIleAla	97
Qy	344	AGTGAGCTGTCACCGCGGCGAGCTGGCTGGCGCGAGCGCTGTGCACGTGTGGATC	403
Db	98	-----TyrThrThrThrArgThrTrpAsnPheGlyGlnIleLeuCysAspIleTrpVal	115
Qy	404	TCCTTCAGCGCGGAGCGCTGTCTGTCTGCCCGCGCGCCCTCGGGAAACGTGGCGCCATC	463
Db	116	SerSerAsp-----IleThrCysCysThrAlaSerIleLeuHisLeuCysValIle	132
Qy	464	GCCCTGGCGCGCAGCGGGCCATCACACGGCACTGCAGACACACGCTCGCACCGCAGC	523
Db	133	AlaLeuAspArgTyrTrpAlaIleThrAspAlaLeuGluTyrSerLysArgArgThrAla	152
Qy	524	CGCGCTCGTTGTCTAATCGCGCTGCCCGGGTGGCGTGGCGGCTCATCCCTCCGCG	583
Db	153	GlyHisAlaGlyAlaMetIleAlaAlaValTrpValIleSerIleCysIleSerIlePro	172
Qy	584	CCGCTGCTCTTTGGCGGGCGAGGTGTGCGACGCTCGGCTCCAGCGCTGCACAGTG	640
Db	173	Pro---LeuPheTrpArgGlnAlaGlnAlaGlnGluMetSerAspCysLeuValAsn	191
Qy	641	AGCCGGAAACCTCCTATGCCGCGCTTCACACCGCGCGGCTTCACCTCCGCTTGGC	700
Db	192	ThrSerGlnIleSerTyrThrIleTyrSerThrCysGlyAlaPheTyrIleProSerVal	211
Qy	701	GTGTGGCGTTTGTCTACCGAAAGATCTACGAGGGGCCAAGTTTGGTTTCGGCGCGC	760
Db	212	LeuLeuIleIleLeuTyrSerArgIleTyrArgAlaIle-----	224
Qy	761	CGGAGAGCTGTGCTGCCGCTTCGCCGCCACCATGTCAGGTGAGGTCCCAAGTAAAGGA	820
Db	225	-----ArgSerArgIleLeuAsnPro	231
Qy	821	CCTGATGAGGCTAAGTGTGTTC---ACGGCACAT-----	853
Db	232	ProSerLeuSerGlyLysArgPheThrThrAlaHisLeuIleThrGlySerAlaGlySer	251
Qy	854	-----TGCMAAGCAACGGTGTCTCTCCAGTGAGC-----	883
Db	252	SerLeuCysSerLeuAsnProSerLeuHisGluGlyHisMetHisProGlySerProLeu	271
Qy	884	-----GGGACATCCTGGCGGGAGCAGAAAG-----	907
Db	272	PhePheAsnHisValArgIleLysLeuAlaAspSerValLeuGluArgLysArgIleSer	291
Qy	908	-----GAGAGCGAGCAGCCATGATGGTGGGAATTCGATTCGATGGCGTGTGTGCTG	958
Db	292	AlaAlaArgGluArgLysAlaThrLysThrLeuGlyIleIleLeuGlyAlaPheIleVal	311
Qy	959	TGCTGGATCCCTTCTTCCTGACGGAATCATCAGCCACATCTCT-----GCCTGCAGC	1012
Db	312	CysTrpLeuProPhePheValSerLeuValLeuProIleCysArgAspSerCysTrp	331
Qy	1013	CTGCCCCCATCTGGAAGAAGCATATTCTGTGGCTTGCTACTCCAAATCTTCTTCAAC	1072
Db	332	IleHisProAlaLeuPheAspPhePheThrTrpLeuGlyTyrLeuAsnSerLeuIleAsn	351
Qy	1073	CCCTGTATTACACAGCTTTTAAACAGAACTACAAACAATGCCTTCAGAGCCTC	1126
Db	352	ProIleIleTyrThrValPheAsnGluAspPheArgGlnAlaPheGlnLysVal	369

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 19:58:47 ; Search time 148.5 Seconds

(without alignments)
4383.767 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgccatggaggccgctagcc.....aagcagagatgaacacagg 1152

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 3172214

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ n2p.model -DEV=xlP
-Q=/cgn2_1/USPRO.spool.P/US09976782/runat 31082004 115949 10296/app query.fasta_1.1351
-DB=A_Geneseq 29Jan04 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosu62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USERS=US09976782 @CGN 1_1 224 @runat 31082004 115949 10296 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq 29Jan04.*

1: Geneseqp1980s.*

2: Geneseqp1990s.*

3: Geneseqp2000s.*

4: Geneseqp2001s.*

5: Geneseqp2002s.*

6: Geneseqp2003as.*

7: Geneseqp2003bs.*

8: Geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1946	90.3	379	5 ABB78809	Abb78809 Human NOV
2	1861.5	86.4	370	5 AAE18654	Aae18654 Human G-p
3	1860.5	86.4	370	4 AAM47211	Aam47211 Human NOV
4	1857.5	86.2	370	4 AAM47212	Aam47212 Human NOV
5	1856.5	86.2	372	5 AAE15638	Aae15638 Human G-p
6	1819.5	84.5	380	5 ABG60235	ABG60235 Human hyd
7	1453.5	67.5	370	2 AAR58686	Aar58686 Rat MR22
8	1453.5	67.5	370	7 ADE56994	Ade56994 Rat Prote
9	1452.5	67.4	370	2 AAR57066	Aar57066 Murine se
10	1452.5	67.4	370	6 AAE31673	Aae31673 Mouse 5-H

11	1296.5	60.2	606	7 ADC86157	Adc86157 Human GPC
12	1063.5	49.4	357	5 AAU79252	Aau79252 Human 5-h
13	1063.5	49.4	357	6 ABP81980	Abp81980 Human 5-H
14	1063.5	49.4	357	7 AAE38595	Aae38595 Human 5-H
15	1063.5	49.4	357	7 ADE56996	Ade56996 Human Pro
16	1058.5	49.1	357	4 ABB56327	Abb56327 Non-endog
17	1053.5	48.9	357	5 AAU79253	Aau79253 Human 5-h
18	1022.5	47.5	357	5 AAR45848	Aar45848 Human 5HT
19	1022.5	47.5	357	5 AAU75167	Aau75167 Mouse 5-H
20	1017.5	47.2	357	2 AAR45847	Aar45847 Murine 5H
21	1014.5	47.1	357	2 AAR58685	Aar58685 Rat REC17
22	647.5	30.1	212	4 AAG62846	Aag62846 Amino aci
23	596	27.7	111	2 AAR57067	Aar57067 Human ser
24	501	23.3	99	4 AAG62845	Aag62845 Amino aci
25	448	20.8	390	2 AAY28304	Aay28304 Serotonin
26	448	20.8	390	4 AAE00911	Aae00911 Human 5-h
27	448	20.8	390	6 ABG73850	Abg73850 Human ser
28	448	20.8	390	6 ABP81761	Abp81761 Human 5-H
29	448	20.8	390	7 ADE40461	Ade40461 Human ser
30	447	20.8	390	2 AAR43060	Aar43060 Human 5-H
31	447	20.8	390	4 ABB56317	Abb56317 Non-endog
32	444	20.6	90	6 ABP72590	Abp72590 Human LP3
33	441.5	20.5	564	4 ABB58411	Abb58411 Drosophil
34	441.5	20.5	564	7 AAE38198	Aae38198 Fruit fly
35	441	20.5	377	6 ABG74161	Abg74161 Dog 5-Ht
36	438	20.3	386	5 ABB57292	Abb57292 Mouse isc
37	431.5	20.0	376	2 AAR15137	Aar15137 Human ser
38	430	20.0	377	2 AAY28303	Aay28303 Serotonin
39	430	20.0	377	4 AAB47185	Aab47185 5-hydroxy
40	430	20.0	377	6 ABG73849	Abg73849 Human ser
41	430	20.0	377	6 ABP81762	Abp81762 Human 5-H
42	430	20.0	377	7 ADE58162	Ade58162 Human pro
43	429	19.9	377	4 ABB56318	Abb56318 Non-endog
44	428	19.9	422	4 AAB70249	Aab70249 HTR1A pro
45	428	19.9	451	5 AABG75677	Abg75677 Human 5HT

ALIGNMENTS

RESULT 1

ABB78809

ID ABB78809 standard; protein; 379 AA.

XX AC ABB78809;

XX DT 29-JUL-2002 (first entry)

XX DE Human NOV5 protein sequence SEQ ID NO:16.

XX DE Human NOV5 protein sequence SEQ ID NO:16.

XX KW Human; NOVX; cytostatic; antiarteriosclerotic; cardiovascular; lymphoma;

XX KW anti-diabetic; immunosuppressive; neuroprotective; gene therapy; cancer;

XX KW cardiomyopathy; atherosclerosis; cell signal processing; diabetes; AIDS;

XX KW metabolic pathway modulation; neoplastic; neurological disorder; asthma;

XX KW adenocarcinoma; prostate cancer; uterus cancer; immune response;

XX KW Crohn's disease; multiple sclerosis; Graft versus host disease;

XX KW chromosome 2.

XX OS Homo sapiens.

XX PN WC2002030974-A2.

XX PD 18-APR-2002.

XX PF 12-OCT-2001; 2001WO-US031922.

XX PR 12-OCT-2000; 2000US-0240113P.

XX PR 16-OCT-2000; 2000US-0240625P.

XX PR 16-OCT-2000; 2000US-0240637P.

XX PR 16-OCT-2000; 2000US-0240648P.

XX PR 16-OCT-2000; 2000US-0240662P.

XX PR 16-OCT-2000; 2000US-0240669P.

XX PR 16-OCT-2000; 2000US-0240703P.

XX PR 16-OCT-2000; 2000US-0240732P.

PR 16-OCT-2000; 2000US-0241190P.
 PR 18-JAN-2001; 2001US-0262455P.
 XX (CURA-) CUPAGEN CORP.
 PA (MILL/) MILLET I.
 XX
 PI Grosse WM, Alsobrook JP, Lopley DM, Burgess CE, Mishra V;
 PI Kekuda R, Li L, Padigaru M, Shimkets RA, Zerhusen BD, Spytke KA;
 PI Edinger S, Gerlach V, Macdougall J, Stone D, Gunther E, Ellerman K;
 XX
 XX WPI; 2002-444172/47.
 DR N-PSDB; AEN86917.
 XX
 PT New NOVX polypeptides and polynucleotides, useful for treating or
 PT preventing a NOVX-associated disorder or a pathological state in a
 PT subject, particularly a human, e.g. cardiomyopathy, atherosclerosis,
 PT cancer or diabetes.
 XX
 XX Claim 1; Page 46; 227pp; English.
 XX
 CC The present invention describes novel human proteins designated NOVX
 CC (where X is 1, 2a, 2b, 2c, 2d, 3, 4, 5, 6a, 6b, 7, 8, or 9). NOV1 is a
 CC tyrosine-protein kinase 6-like protein; NOV2a-d are keratin 4-like
 CC proteins; NOV3 is a collagen-like protein; NOV4 is a cystatin B-like
 CC protein; NOV5 is a serotonin receptor-like protein; NOV6a and NOV65v are
 CC cold inducible glycoprotein 30-like proteins; NOV7 is a matrilin-2-like
 CC protein; NOV8 is a leukocyte surface antigen (CD53)-like protein; and
 CC NOV9 is a tyrosine kinase-like protein. NOVX sequences have cytostatic,
 CC antiatherosclerotic, cardiovascular, antidiabetic, immunosuppressive and
 CC neuroprotective activities, and can be used in gene therapy. The NOVX
 CC sequences can be used in therapeutics, particularly for treating,
 CC preventing or alleviating a NOVX-associated disorder or a pathological
 CC state in a subject, particularly a human. These disorders include
 CC cardiomyopathy, atherosclerosis, a disorder related to cell signal
 CC processing and metabolic pathway modulation or diabetes. The NOVX
 CC sequences are also useful for determining the presence of or
 CC predisposition to a disease associated with altered levels of NOVX
 CC polypeptide or nucleic acid, particularly cancer. The NOVX sequences are
 CC especially useful in therapeutic or prophylactic applications for
 CC neoplastic or neurological disorders, and in the treatment of
 CC adenocarcinoma, lymphoma, prostate cancer, uterus cancer, immune
 CC response, AIDS, asthma, Crohn's disease, multiple sclerosis or Graft
 CC versus host disease. The present sequence represents the human NOV5
 CC protein from the present invention. NOV5 is located to chromosome 2
 XX
 SQ Sequence 379 AA;

Alignment Scores:
 Pred. No.: 4,89e-143 Length: 379
 Score: 1946.00 Matches: 379
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 90.34% Indels: 0
 DB: 5 Gaps: 0

US-09-976-782-15 (1-1152) x ABB78809 (1-379)

QY 5 ATGGAGCGCGTAGCCTTTCAGTGGCCACCGCGGGTTCCTTGCCTGGGACCCGAG 64
 Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
 QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGGGATACCTCGTTCCGACCCCGAGCGGC 124
 Db 21 ThrSerSerGlyProGlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 40
 QY 125 GCGCTCTCGCGGCGGAGGCGCCCTTCTCTGCTTTCACGGTCTCGTGTGAGCGCTG 184
 Db 41 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 60
 QY 185 CTAGTGTCTGATCGCTGCACCTTTCCTGTGGAACCTGCTGCTCGGTCCACCATCCCG 244
 Db 61 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValProValThrIlePro 80

QY 245 CGGTCCTCGTCCCTTCCACCGCGTCCGCGATACCTTGGTGGCTCGACGCGCTCTCGGAC 304
 Db 81 ArgValArgAlaPheHisArgValProHisLeuValAlaSerThrAlaValSerAsp 100
 QY 305 GAATAGTGGCAGCGCTCGGATGACCCAGCGCTGGCGAGTGAAGTGTCTGACCGGGGA 364
 Db 101 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 120
 QY 365 GGTCCGCTGCTGGCGCGGAGCGCTGGCCACGCTGTGATCTCTTCCGACGCGCGAGCCTGT 424
 Db 121 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAlaGlyAlaCys 140
 QY 425 CTGTCTCGTCCCGCGCGCTCGGGAACCTGGCGGCATCGCCCTGGCGCGACGCGGCC 484
 Db 141 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 160
 QY 485 ATCACACGCGACCTCGACACACGCTGGCGACCGCGCGCTCGTGTCTCATGATC 544
 Db 161 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 180
 QY 545 GCGCTCGCGCGGCTCGGCTCGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGGC 604
 Db 181 AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 200
 QY 605 GAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGGTGAGCGCGGAACCTCTTATGCGGCC 664
 Db 201 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 220
 QY 665 TTCTCCACCGCGCGCGCTTCCACCTGGCTGGCGTGGCGTGGCTGCTGTCTACCGGAG 724
 Db 221 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys 240
 QY 725 ATCTACGAGGCGGCAAGTTTCGTTTCGCGCGCGCGAGAGCTGTGCTGCGTGTGCGG 784
 Db 241 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuLeuProLeuPro 260
 QY 785 GCCACCATCGAGGTGAGGTCCAAAGTAAGGAAGCACCTGTAGGTGAGGTGAGTGTTC 844
 Db 261 AlaThrMetGlnValArgSerLysValLysGluAlaProAspGluAlaGluValValPhe 280
 QY 845 ACGGCACATTGCAAGCAACGCTGCTTCCAGGTGAGCGGGGACTCTTGGCGGAGCAG 904
 Db 281 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 300
 QY 905 AAGGAGAGCGAGCAGCATGATGCTGGGAATTCCTGATTGGCGTGTTCGTGCTGCTCGG 964
 Db 301 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 320
 QY 965 ATCCCTCTTCTTCTCGACGGAACCTCATCAGCCCACTCTGTGCTCGAGCTGCCCCCAATC 1024
 Db 321 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 340
 QY 1025 TGGAAAAGCATATTTCTGTGGCTTGGCTACTCCAATTTCTTCTTCAACCCCTGATTAC 1084
 Db 341 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 360
 QY 1085 ACAGCTTTTAAACAAGAACTACAAACAATGCTTCAAGAGCTCTTACTAAGCAGAGA 1141
 Db 361 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 379

RESULT 2
 AAE18654
 ID AAE18654 standard; protein; 370 AA.
 XX
 AC AAE18654;
 XX
 DT 17-MAY-2002 (first entry)
 XX
 DE Human G-protein coupled receptor (GREC-15).
 XX
 KW Human; G-protein coupled receptor; GREC-15; cell proliferative disorder;
 KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;
 KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;
 KW

QY 665 TTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGTGGCGTGTCTACCGGAG 724
 Db 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys 235
 QY 725 ATCTACGAGCGCGCAAGCTTTCTGTCGCGCGCGCGAGAGCTGTCTGCGGTGCGG 784
 Db 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
 QY 785 GCCACCATGTCAGGTGAGTCCAAAGTAAAGGAGACCTGATGAGGCTGAAGTGTTC 844
 Db 256 AlaThr-----SerLysValLysGluAlaProAspGluAlaGluValValPhe 271
 QY 845 ACGGCACATGTCAAAGCAACGCTTCCAGGTGAGCGGACTCTCTGGCGGAGCAG 904
 Db 272 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGGAGCGGAGCAGCATCATGTGGGAATTCGATGTCGTCGTGTGTCGTGG 964
 Db 292 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTTCTCTGACGGAATCATCAGCCCACTCTGTGCTGTCAGCTGCCCGCCATC 1024
 Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
 QY 1025 TGGAAAGCATATTTCTGTGCTGGCTTGGCTACTCCAATCTTTCTCAACCCCTGATTTC 1084
 Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
 QY 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCCTCTTTACTTAAGCAGAGA 1141
 Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 4

AA047212
 ID AA047212 standard; protein; 370 AA.

XX AC AA047212;
 DT 12-FEB-2002 (first entry)
 XX XX Human NOV5b protein.
 KW Human; NOV5b; metabolic disorder; neurodegenerative disorder;
 KW immune disorder; haematopoietic disorder; developmental disease; cancer;
 KW retinal disease; feeding disorder; vaccine; infection; gene therapy;
 KW neurological disorder; psychotic disorder; G-protein coupled receptor;
 KW analgesic; antidiabetic; viricide; neuroprotective; nootropic;
 KW anesthetic; antidepressant; antimigraine; anticonvulsant; neuroleptic;
 KW antipsychotic; antiallergic; antiinflammatory; anorectic; antiarthritic;
 KW osteoporotic; antiatherosclerotic; antibacterial; fungicide;
 KW osteoplastic; protozoacide; antiulcer; hypertensive; hypotensive;
 KW antiinfertility; vulnular; nephrotropic; antilipemic;
 KW serotonin receptor; chromosome 2.

OS Homo sapiens.

PH Key Location/Qualifiers
 FT Peptide 1..24
 FT Protein /label= signal_peptide
 FT Protein 25..370
 FT Protein /label= mature_NOV5b

DN W0200174851-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US010039.

XX 30-MAR-2000; 2000US-0193205P.

XX 30-MAR-2000; 2000US-0193339P.

XX 05-APR-2000; 2000US-0195343P.

XX 06-APR-2000; 2000US-0195005P.

PR 06-APR-2000; 2000US-0195088P.
 PR 10-APR-2000; 2000US-0195792P.
 PR 11-APR-2000; 2000US-0196556P.
 PR 13-APR-2000; 2000US-0197081P.
 PR 14-APR-2000; 2000US-0197087P.
 PR 14-APR-2000; 2000US-0197525P.
 PR 29-MAR-2001; 2001US-00823187.
 XX (CURA-) CURAGEN CORP.
 XX Majumder K, Spaderna SK, Taupier RJ, Padigaru M, Burgess CE;
 PI Shimkets RA, Spytek KA, Liu X, Patturajan M, Gusev VV;
 XX WPI; 2001-626379/72.
 DR N-PSDB; ABA01985.
 XX
 PT New G protein-coupled receptor related polypeptides and polynucleotides
 PT for diagnosis, prevention and treatment of metabolic, neurodegenerative,
 PT retinal, immune, hematopoietic disorders, diabetes, obesity and
 PT infections.
 XX

PS Claim 1; Page 47; 194pp; English.

CC The present invention provides the protein and coding sequences of novel
 CC human G-protein coupled receptors, designated NOV1, NOV2, NOV3, NOV4,
 CC NOV5a, NOV5b, NOV6a, NOV6b, NOV7, NOV8a, NOV8b, NOV9 and NOV10. These can
 CC be used in the treatment of NOVX related diseases, including cancer,
 CC metabolic, neurodegenerative, immune, haematopoietic, developmental,
 CC retinal, feeding, neurological and psychotic diseases and disorders and
 CC infections. The present sequence is the NOV5b protein, the gene for which
 CC is found on chromosome 2. The NOV5b protein shares homology with
 CC serotonin receptors
 XX

SQ Sequence 370 AA;

Alignment Scores:

Pred. No.: 3,82e-136 Length: 370
 Score: 1857.50 Matches: 369
 Percent Similarity: 97.36% Conservative: 0
 Best Local Similarity: 97.36% Mismatches: 1
 Query Match: 86.23% Indels: 9
 DB: 4 Gaps: 3

US-09-976-782-15 (1-1152) x AA047212 (1-370)

QY 5 ATGAGAGCGCGTACCTTTAGTGGCCACCGCGCGGTGCCCTTGCCTGGGACCGAG 64
 Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20
 QY 65 ACCAGCAGCGGACCGCGGACCCCAAGCCGAGAGGGATCTCGGTTCGACCCCGAGCGGC 124
 Db 21 ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 38
 QY 125 GCCGTCTGCGGCGCGAGGCGCCCTTCTCTCTTCTACGTCCTGGTGGTGGACGTG 184
 Db 39 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 58
 QY 185 CTAGTGTGTGATCGCTGCCACTTCTCTGTGGAACCTGCTGGTTCGGTCCGTCACATCCG 244
 Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro 78
 QY 245 CGGTCGCTGCTTCCACCGCGTCCGCGTAACTTGGTGGCTCGACGCGCGTCTCGGAC 304
 Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAAGTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGGAGTGTGACCGCGCGA 364
 Db 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
 QY 365 CGTCGCTGTGGCGCGAGCCTGTGCCAGTGTGGATCTCTTCGACGCGCGAGCCTGT 424
 Db 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheAspAla----- 135

```
QY 425 CTGTGCTGCGCGCGCTCGGGAACGTGGCGGCATCGCGCTGGCGCGACGGGCGC 484
Db |||||||
QY 136 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaAlaLeuGlyArgAspGlyAla 155
Db |||||||
QY 485 ATCACACGCGACCTCGAGCACACGCTGGCGCACCGCGCGCTGCTGCTCATGATC 544
Db |||||||
QY 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 175
Db |||||||
QY 545 GCGCTGCGCGCGGTCGCGCTGCGCTCATCGCGCTGCGCGCGCTGCTTGGCGCGGCGC 604
Db |||||||
QY 176 AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 195
Db |||||||
QY 605 GAGGTGTCAGCGCTCGGCTCGAGCGCTGCCAGGTGAGCGCGGAACCTCTATGCGCGC 664
Db |||||||
QY 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 215
Db |||||||
QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCGCTGGCGTGGTGGTGGTGGTGGTGGTGG 724
Db |||||||
QY 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValAlaProPheValTyrArgLys 235
Db |||||||
QY 725 ATCTACGAGCGCGCAAGTTGTTTCGCGCGCGCGCGAGAGCTGCTGCGGTGCGG 784
Db |||||||
QY 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgAlaValLeuProLeuPro 255
Db |||||||
QY 785 GCCACCATCGAGTCCCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGGTGTTC 844
Db |||||||
QY 256 AlaThrMetGln-----ValLysGluAlaProAspGluAlaGluValPhe 271
Db |||||||
QY 845 ACGGCACATGCAAGCAAGGTGCTTCCAGGTGAGCGGGGACTCTGCGCGGAGCAG 904
Db |||||||
QY 272 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 291
Db |||||||
QY 905 AAGGAGCGCGAGCGCATGATGTGGGAATCTGATTGGCGTGTGCTGTGCTGCG 964
Db |||||||
QY 292 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
Db |||||||
QY 965 ATCCCTCTTCTCTGACGGAATCATACGCCCACTGTGCTGCTGAGCGCTGCGCCCATC 1024
Db |||||||
QY 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProIle 331
Db |||||||
QY 1025 TCGAAAGCATATTTCTGGGTTGGCTACTCAATCTTCTTCAACCCCTGATTATC 1084
Db |||||||
QY 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
Db |||||||
QY 1085 ACAGCTTTAAACAGCACTACACATGCTTCAGAGCTCTTTACTAAGCAGAGA 1141
Db |||||||
QY 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
Db |||||||

RESULT 5
AAE15638
ID AAE15638 standard; protein; 372 AA.
XX
AC AAE15638;
XX
DT 12-MAR-2002 (first entry)
XX
DE Human G-protein coupled receptor-8 (GCRC-8) protein.
XX
KW Human; G-protein coupled receptor-8; GCRC-8; therapy; cancer; stroke;
KW cell proliferative disorder; neurological; epilepsy; Parkinson's disease;
KW Alzheimer's disease; inflammation; thyroiditis; haemolytic anaemia; AIDS;
KW Acquired Immune Deficiency Syndrome; dementia; nontropic; cholelithiasis;
KW multiple sclerosis; atherosclerosis; angina pectoris; gastroenteritis;
KW diabetes; ulcer; viral infection; immunosuppressive.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..65
FT FT /label= Signal_peptide
FT FT 48..68
FT FT /label= Transmembrane_domain
FT FT 66..372
```

```
FT Domain /label= Mature_GCRC_8_protein
FT 300..318
FT FT /label= Transmembrane_domain
XX
```

PN WO200198351-A2.

XX PD 27-DEC-2001.

XX PF 15-JUN-2001; 2001WO-US019275.

XX PR 16-JUN-2000; 2000US-0212483P.

XX PR 22-JUN-2000; 2000US-0213954P.

XX PR 29-JUN-2000; 2000US-0215209P.

XX PR 07-JUL-2000; 2000US-0216595P.

XX PR 14-JUL-2000; 2000US-0218936P.

XX PR 19-JUL-2000; 2000US-0219154P.

XX PR 21-JUL-2000; 2000US-0220141P.

XX PA (INCY-) INCYTE GENOMICS INC.

XX PI Lal P. Baughn MR, Hafalia AJA, Nguyen DB, Gandhi AR, Kallick DA;

XX PI Griffin JA, Yue H, Khan FA, Patterson C, Lu DAM, Tribouley CM;

XX PI Lu Y, Wallia NK, Graul R, Yao MG, Yang J, Ramkumar J, Au-Young J;

XX PI Elliott VS, Hernandez R, Walsh RT, Borowsky ML, Thornton M, He A;

XX WPI; 2002-075627/10.

XX DR N-PSDB; AAD24963.

XX PT Isolated human G-protein coupled receptor polypeptides and the use of

XX PT these sequences in the diagnosis, treatment and prevention of diseases

XX PT and in the assessment of exogenous compounds on the expression of the

XX PT receptors.

XX PS Claim 1; Page 121; 143pp; English.

XX CC The invention relates to isolated human G-protein coupled receptor

XX CC (GCRC) polypeptides and their biologically active fragments. GCRC and

XX CC protein is useful in treating a disease or condition associated with an

XX CC increase or decrease in expression of functional GCRC. The GCRC's are

XX CC useful in the diagnosis, treatment and prevention of cell proliferative

XX CC disorders (cancer, leukemia, melanoma); neurological disorders (stroke,

XX CC epilepsy, Parkinson's disease, dementia, Alzheimer's disease); autoimmune

XX CC inflammatory disorder (thyroiditis, haemolytic anaemia, AIDS, multiple

XX CC sclerosis); cardiovascular disorder (atherosclerosis, angina pectoris),

XX CC gastrointestinal disorder (ulcer, cholelithiasis, gastroenteritis),

XX CC metabolic disorders (diabetes); viral infections (herpes virus) and in

XX CC the assessment of the effects of exogenous compounds on the expression of

XX CC the nucleic acid and amino acid sequences. The present sequence is human

XX CC GCRC-8 protein

XX SQ Sequence 372 AA;

Alignment Scores:

Pred. No.: 4,58e-136 Length: 372

Score: 1856.50 Matches: 369

Percent Similarity: 97.36% Conservative: 0

Best Local Similarity: 97.36% Mismatches: 3

Query Match: 86.19% Indels: 7

DB: 5 Gaps: 3

US-09-976-782-15 (1-1152) x AAE15638 (1-372)

QY 5 ATGGAGCGCGCTAGCCTTTTCAGTGGCCACCGCGCGTTCCTTCCTGGGACCCGAG 64

Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20

QY 65 ACCAGCAGCGACCCCGGACCCCAAGCCGAGGAGGATCTCGTTTCGACCCCGACCGCG 124

Db 21 ThrSerSer-----GlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 38

QY 125 GCGGTCTCTCGCGCGCGGCGCGCGCTTCTCTCTTCACGCTCTCTGGTGGTGAGCGCTG 184

Db 39 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValThrLeu 58

Qy 185 CTAGTGTGCTGATCGTGGCCACTTCTCTGGAACCTGCTGGTTCGGTCCACCATCCG 244
Db |||||
Qy 59 LeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro 78
Db |||||
Qy 245 CGGTCCTGCTCCACCGCGTGGCGGATACCTTGGTGGCTCGACGGCGCTTCGAC 304
Db |||||
Qy 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
Qy 305 GAACCTAGTGGCAGCGTGGCGATGCCAGCGCTGGCGAGTGGAGTGTCTGACCGGGCGA 364
Db |||||
Qy 99 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 118
Qy 365 CGTGGCTGTGCGCGCGAGCCTGTGCCACCTGTGGATCTCTTCGACGCGCGAGCGCTGT 424
Db |||||
Qy 119 ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheHisVal----- 135
Qy 425 CTGTGCTGCCCCCGCGCTCGGGAACGTGGCGGCTATCGCCCTGGCGCGGACGGGGCC 484
Db |||||
Qy 136 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 155
Qy 485 ATCACAGGACCGCGAGCAGCACGCTGGCGACCGCGCGCGCTCGTCTGCTCATGATC 544
Db |||||
Qy 156 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 175
Qy 545 CGCTCGCCCGCGGTGCGCTCGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGCGGGC 604
Db |||||
Qy 176 AlaLeuThrArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 195
Qy 605 GAGGTGTGCGAGCTCGGCTCCAGCGTCCAGGTGAGCGCGGAAACCTCTCTATGCGGCC 664
Db |||||
Qy 196 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla 215
Qy 665 TTCTCCACCGCGCGCTCCACCTCGCGTGGCGTGGCGTGGCTGTCTGTCTACCGAAG 724
Db |||||
Qy 216 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValProPheValTyrArgLys 235
Qy 725 ATCTACGAGCGCGCAAGTTCGTTCGCGCGCGCGGAGCTGTGCTGCGCTGCGG 784
Db |||||
Qy 236 IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro 255
Qy 785 GCACCATGAGGTGAGGTCAAGTAAAGAAAGCACTGATGAGGTGAGGTGAGTGTTC 844
Db |||||
Qy 256 AlaThrMetGlnVal-----LysValLysGluAlaProAspGluAlaGluValValPhe 273
Qy 845 ACGGCATTTGCAAGCAACGTTCTCTCCAGGTGAGCGGGACTCTGCGCGGAGCAG 904
Db |||||
Qy 274 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 293
Qy 905 AAGGAGAGCGGAGCAGCATGATGGTGGGAATCTGATTGCGGTGTGTGTGCTGTGCG 964
Db |||||
Qy 294 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 313
Qy 965 ATCCCTCTCTCTGAGGGAACCTCATGAGCCCACTGTGCTGCTGCGAGCTGCGCCCATC 1024
Db |||||
Qy 314 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProIle 333
Qy 1025 TGAAAAGCATATTTCTGCTGCTGCTGCTCCCAATCTTCTTCAACCCCTGATTAC 1084
Db |||||
Qy 334 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 353
Qy 1085 ACAGCTTTTAAAGAACTCAACAATGCTTCAAGAGCTCTTTACTAAGCAGAGA 1141
Db |||||
Qy 354 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 372
RESULT 6
ABG60235
ID ABG60235 standard; protein; 380 AA.
XX AC ABG60235;
XX DT 30-JUL-2002 (first entry)
XX

DE XX Human hydroxytryptamine receptor-like protein NOV8.
KW Human; NOVX; developmental disorder; endocrine disorder;
KW vascular disorder; infectious disease; anorexia; cancer; stroke;
KW neurodegenerative disorder; Alzheimer's disease; acute brain injury;
KW central nervous system disorder; depression; lung disorder;
KW reproductive disorder; tissue disorder; thrombocytopaenia; migraine;
KW angiogenesis; asthma; X-linked severe combined immunodeficiency;
KW inflammation; autoimmune disorder; immune disorder; blood disorder;
KW haematopoietic disorder; gastrointestinal disease; respiratory disorder;
KW hepatitis; fertility; hypertension; arteriosclerosis; ischaemia;
KW rheumatoid arthritis; Grave's disease; wound healing.
XX OS Homo sapiens.
XX WO200224733-A2.
XX 28-MAR-2002.
XX 17-SEP-2001; 2001WO-US029115.
XX 15-SEP-2000; 2000US-0232675P.
XX 15-SEP-2000; 2000US-0232676P.
XX 15-SEP-2000; 2000US-0232679P.
XX 18-SEP-2000; 2000US-0233382P.
XX 18-SEP-2000; 2000US-0233402P.
XX 19-SEP-2000; 2000US-0233521P.
XX 19-SEP-2000; 2000US-0233522P.
XX 19-SEP-2000; 2000US-0233801P.
XX 20-SEP-2000; 2000US-0233960P.
XX 06-OCT-2000; 2000US-0238398P.
XX 13-OCT-2000; 2000US-0240284P.
XX 13-OCT-2000; 2000US-0240498P.
XX 11-JAN-2001; 2001US-0260973P.
XX 26-JAN-2001; 2001US-0264274P.
XX 09-MAR-2001; 2001US-0274862P.
XX (CURA-) CURAGEN CORP.
XX Mishra VS, Syptek KA, Taupier RJ, Vernet CAM, Colman SD;
XX Gorman L, Tchernev VT, Malyankar UM, Shenoy S, Tchernev VT;
XX Padigaru M, Patturajan M, Burgess CE, Smithson G, Millet I;
XX Peyman JA, Stone D, Gunther E, Ellerman K;
XX WPI: 2002-383182/41.
XX N-PSDB; ABK1923.
XX New cytoplasmic, nuclear, membrane bound and secreted NOVX polypeptides,
XX useful for treating cancers and tumors, lung disorders, hematopoietic
XX disorders, autoimmune diseases and immune disorders.
XX Claim 1 ; Page 60; 210pp; English.
XX The invention relates to an isolated NOVX polypeptide selected from
XX NOV1a, NOV1b, NOV1c, NOV2a, NOV2b, NOV2c, NOV3a, NOV3b, NOV4a, NOV4b,
XX NOV5a, NOV5b or NOV6-NOV9 polypeptides, their mature form or variant.
XX Also included are a nucleic acid encoding a NOVX protein or variant; a
XX vector comprising the nucleic acid; a cell comprising the vector; an anti
XX -NOVX antibody; and identifying agents that modulate the expression or
XX activity of NOVX. NOVX, the nucleic acid, antibody and modulators are
XX useful in the diagnosis, treatment or prevention of developmental
XX disorders, endocrine disorders, vascular disorders, infectious disease,
XX anorexia, cancer, neurodegenerative disorders (e.g. Alzheimer's disease,
XX Parkinson's disease, Huntington's disease, multiple sclerosis and
XX amyotrophic lateral sclerosis), acute brain injury (e.g. stroke, head
XX injury and cerebral palsy), central nervous system disorders (e.g.
XX depression, epilepsy and schizophrenia), lung disorders, reproductive
XX disorders, disorders affecting carbohydrate metabolism (e.g.
XX galactosaemia and hereditary fructose intolerance), tissue disorders
XX (e.g. Wiskott-Aldrich syndrome, thrombocytopaenia, night blindness and
XX Pick's disease), disorders linked to abnormal angiogenesis, asthma,
XX azoospermia, learning disabilities, facial dysmorphism, autoimmune
XX encephalomyelitis, X-linked severe combined immunodeficiency, seizures,

CC migraines, inflammation, autoimmune disorders, disorders affecting sleep,
 CC appetite, thermoregulation, pain, perception, hormone secretion and
 CC sexual behaviour, immune disorders, haematopoietic disorders or other
 CC disorders related to cell signal processing and metabolic pathway
 CC modulation, gastrointestinal diseases, respiratory disorders, blood
 CC disorders, hepatitis, trauma, regeneration, viral, bacterial or parasitic
 CC infections, hyper- or hypo-thyroidism, endometriosis, fertility,
 CC hypertension, arteriosclerosis, ischaemia, haemolytic anaemia, Werner
 CC syndrome, rheumatoid arthritis, Grave's disease, wound healing, X-linked
 CC mental retardation, psychotic and neurological disorders and neuronal
 CC degeneration. The present sequence represents a NOVX protein
 XX
 SQ Sequence 380 AA;

Alignment Scores:
 Pred. No.: 3,51e-133 Length: 380
 Score: 1819.50 Matches: 367
 Percent Similarity: 93.62% Conservative: 0
 Best Local Similarity: 93.62% Mismatches: 0
 Query Match: 84.47% Indels: 25
 DB: 5 Gaps: 5

US-09-976-782-15 (1-1152) x ABG60235 (1-380)

QY 5 ATGGAGGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTTCCTTCCTCGCCGACCCGAG 64
 DB 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAla-----ProGlu 18
 QY 65 ACCAGCAGC-----GGACCGGGACC 85
 DB 19 ThrSerSerProAlaLeuProLeuProTrpAspProArgProAlaAlaGlyProGlyThr 38
 QY 86 CCAAGCCGAGAGGATACCTGGTTCGACCCGCGCGCGGTCTCTGCGCGCGCGAGGG 145
 DB 39 ProSerProArgGlyLeuGlySerThrProSerGlyAlaValLeuProGlyArgGly 58
 QY 146 CGGCCCTCTCTGCTTACGCTGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 205
 DB 59 ProProPheSerValPheThrValLeuValValThrLeuLeuValLeuLeuAlaAla 78
 QY 206 ACTTTCTCTGGAACCTGCTGTTCCGTCACATCCCGCGCGGTCTGCTTCACCGC 265
 DB 79 ThrPheLeuTrpAsnLeuLeuValProValThrIleProArgValArgAlaPheHisArg 98
 QY 266 GTGCGCATAACTGTTGCTGCTGCGCGCGGTCTGCGAGAACTAGTGGCAGCGTGGGG 325
 DB 99 ValProHisAsnLeuValAlaSerThrAlaValSerAspGluLeuValAlaAlaLeuAla 118
 QY 326 ATGCCACCGAGCTGCGGAGTGGTGTGACCGCGCGGACGTGCGGTGGCGCGGAGC 385
 DB 119 MetProProSerLeuAlaSerGluLeuSerThrGlyArgArgArgLeuLeuGlyArg--- 137
 QY 386 CTGTGCCAGTGTGATCTCTTTCGACCGCGGAGCCTGTCTGTGTGCGCCCGCGCGCTC 445
 DB 138 -----HisValTrpIleSerPheAspAla-----LeuCysCysProAlaGlyLeu 152
 QY 446 GGGACGTCGCGCGCATCCCTGGCGCGCGGCGCATCACAGCGCACCTTCAGAGCAC 505
 DB 153 GlyAsnValAlaAlaAlaLeuGlyArgAspGlyAlaIleThrArgHisLeuGlnHis 172
 QY 506 AGCTGTCGACCGCGCGCGCTGTTGCTCATGTCGCTGCGCTGCGCGCGGTGGCGTGC 565
 DB 173 ThrLeuArgThrArgSerArgAlaSerLeuLeuMetIleAlaLeuAlaArgValProSer 192
 QY 566 GGGCTCATCGCTCGCGCGCTGCTCTTTGGCGGGGCGAGTGGACGCTCGGCTC 625
 DB 193 AlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGlyGluValCysAspAlaArgLeu 212
 QY 626 CAGCGCTGCGAGTGGCGGAAACCTCTATGCGCGCTTCTCCACCGCGCGCGCTTC 685
 DB 213 GlnArgCysGlnValSerArgGluProSerTyAlaAlaPheSerThrArgGlyAlaPhe 232
 QY 686 CACCTGCGCGCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 745

DB 233 HisLeuProLeuGlyValValProPheValTyArgLysIleTyGluAlaAlaLysPhe 252
 QY 746 GCTTTCGCGCCCGCGCGAGAGCTGTGCTGCGGTTCGCGCCACCATCATGAGTGGAGTCC 805
 DB 253 ArgPheGlyArgArgArgAlaValLeuProLeuProAlaThrMetGln----- 269
 QY 806 AAGGTAAAGGAAGCACCTGATGAGGCTCAAGTGGTGTTCACGGCACATTCGAAAGCAAG 865
 DB 270 ---ValLysGluAlaProAspGluAlaGluValValPheThrAlaHisCysLysAlaThr 288
 QY 866 GTGTCTCTTCAGGTGAGCGGGACTCTCTGCGGGAGCAGAGGAGGAGGAGGAGGAGGAGG 925
 DB 289 ValSerPheGlnValSerGlyAspSerTrpArgGluGlnLysGluArgAlaMet 308
 QY 926 ATGTGGGAATCTCATTTGGCGGTGTTGTGCTGTGATCCCTCTCTCTGACGGA 985
 DB 309 MetValGlyIleLeuIleGlyValPheValLeuGlyTrpIleProPheLeuThrGlu 328
 QY 986 CTCATCAGCCCACTCTGTGCTGTGAGCTGCGCCCGCATCTGGAAAGCATATTTCTGTGG 1045
 DB 329 LeuIleSerProLeuCysAlaCysSerLeuProProIleTrpLysSerIlePheLeuTrp 348
 QY 1046 CTTGCTACTCAATCTTCTTCAACCCCTGATTTACAGCTTTTAAAGAACTAC 1105
 DB 349 LeuGlyTySerAsnSerPhePheAsnProLeuIleTyThrAlaPheAsnLysAsnTy 368
 QY 1106 AACATGCTCTCAAGAGCTCTTTTACTAAGCAGAGA 1141
 DB 369 AsnAsnAlaPheLysSerLeuPheThrLysGlnArg 380
 RESULT 7
 AAR58686
 ID AAR58686 standard; protein; 370 AA.
 XX AAR58686;
 XX 25-MAR-2003 (revised)
 DT 24-MAY-1995 (first entry)
 XX
 DE Rat MR22 serotonin receptor protein.
 XX
 KW Serotonin; receptor; transmembrane; domain; kinase; phosphorylation;
 KW sensory; motor; behaviour; central nervous system; CNS; superfamily;
 KW G-protein; ligand-gated; ion channel; subfamily; human; rat; amplify;
 KW primer; PCR; amplification; brain; hypothalamus; indolamine; drug;
 KW hypothalamus; therapeutic; neurological; pathology; dementia; insomnia;
 KW Parkinson's disease; eating disorder; anxiety; migraine; headache.
 XX
 OS Rattus rattus.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 5
 FT Domain /label= N-linked glycosylation
 FT Domain 53..78
 FT Domain /label= transmembrane domain
 FT Domain 89..115
 FT Domain /label= transmembrane domain
 FT Domain 126..149
 FT Domain /label= transmembrane domain
 FT Modified-site 163
 FT Modified-site /label= kinase C phosphorylation site
 FT Modified-site 166
 FT Domain /label= kinase C phosphorylation site
 FT Domain 170..196
 FT Domain /label= transmembrane domain
 FT Domain 211..237
 FT Modified-site /label= transmembrane domain
 FT Modified-site 272
 FT Modified-site /label= kinase C phosphorylation site
 FT Modified-site 287
 FT Domain /label= kinase C phosphorylation site
 FT Domain 298..324

AC AAE31673;
 XX 07-MAR-2003 (first entry)
 XX Mouse 5-HT5B receptor protein.
 DE Mouse; 5-hydroxytryptophan; 5-HT5B receptor; depression; transgenic;
 KW transgenic animal; phenotype; pain sensitivity; neurological disease;
 KW gene therapy; pharmacological; neuropsychological disease; analgesic;
 KW psychotic illness; receptor.
 XX Mus sp.
 XX WO200279443-A2.
 PN 10-OCT-2002.
 PD
 XX 29-MAR-2002; 2002WO-US009853.
 XX 29-MAR-2001; 2001US-0280553P.
 PR 21-DEC-2001; 2001US-0342472P.
 PR 28-MAR-2002; 2002US-00109532.
 XX (DELT-) DELTAGEN INC.
 PA
 XX Allen KD;
 PI
 XX WPI; 2003-067450/06.
 DR N-PSDB; AAD48747.
 DR
 XX Novel transgenic mouse useful for identifying a potential therapeutic
 PT agent for the treatment of depression or pain, comprises disruption in a
 PT 5-hydroxytryptophan 5B receptor gene.
 XX Disclosure; Fig 2; 60pp; English.
 PS
 XX The present invention relates to transgenic animals, compositions and
 CC methods relating to the characterisation of gene function. The invention
 CC also relates to transgenic mice comprising mutations in 5-hydroxy-
 CC tryptophan (5-HT)5B receptor gene. The transgenic mice are useful for
 CC identifying an agent that modulates the phenotype such as increased
 CC depression or increased pain sensitivity. They are also useful for
 CC identifying potential therapeutic agents for the treatment of pain or
 CC depression and for evaluating a potential therapeutic agent capable of
 CC affecting a condition associated with a mutation in a 5-HT5B receptor
 CC gene. Transgenic animals of the invention are also useful for testing the
 CC efficacy of proposed genetic and pharmacological therapies for human
 CC diseases such as neurological, neuropsychological or psychotic illnesses.
 CC The present sequence is mouse 5-HT5B receptor protein
 XX
 SQ Sequence 370 AA;
 Alignment Scores:
 Pred. No.: 1.36e-104 Length: 370
 Score: 1452.50 Matches: 298
 Percent Similarity: 83.64% Conservative: 19
 Best Local Similarity: 78.63% Mismatches: 53
 Query Match: 67.43% Indels: 9
 DB: 6 Gaps: 3
 US-09-976-782-15 (1-1152) x AAE31673 (1-370)
 QY 5 ATGGAGCGCGTAGCCTTTCACTGCGCCACCGCGCGGTGGCTTCCCTCGGACCGGAG 64
 DB 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20
 QY 65 ACCAGCAGCGGACCGGGACCCCAAGCCCGAGAGGGATACCTCGTTTCGACCCCGAGCGGC 124
 DB 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
 QY 125 GCGTCTCTCGCGCGCGAGGGCGCGCTTCTGTCTTCAACCGTCTGTGTGTGTGTGTGTGT 184
 DB 39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58

QY 185 CTAGTGTGCTGATCGTCCCACTTTCTGTGTGAACCTGCTGTTCCGGTCCACCATCCCG 244
 DB 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrIleLeu 78
 QY 245 CCGGTCCGTGCTTCACCGCGTGGCATAACTTGGTGGCTCGACAGCGCGCTCTCGGAC 304
 DB 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
 QY 305 GAATAGTGGCAGCGCTGGCATGCCACCGACCTGGCGAGTGGAGTGTGTGTGTGTGTGT 364
 DB 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
 QY 365 CGTGGCTGCTGGCGCGGAGCTGTGCCACGTGTGATCTCTCTTCGACCGCGAGCGCTGT 424
 DB 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
 QY 425 CTGTGTCGCCCGCGGCTTCGGGAACGTGGCGGCATCGCCCTGGCGCGACGGGGCC 484
 DB 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
 QY 485 ATCACACGGCAGCTGCAGCAGCTGGCGCACCGCGAGCGCGCTGTGTGTGTGTGTGTGT 544
 DB 156 IleThrArgHisLeuGlnTrpThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
 QY 545 GCGTCCGCGCGGTGCGCGCTCATCGCCCTCGCGCGCTGTCTTTGGCGCGGCG 604
 DB 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProleuLeuPheGlyTrpGly 195
 QY 605 GAGGTGTGCGAGCTGCGCTCCAGCGTCCAGGTGAGCGCGGAAACCTCTATGCGCGC 664
 DB 196 GluAlaTrpAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTrpAlaVal 215
 QY 665 TTCTCCACCGCGCGCTTCACCTGGCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 724
 DB 216 PheSerThrCysGlyAlaPheTrpLeuProLeuAlaValLeuPheValTrpTrpLys 235
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGCGCGCGGAGAGCTGTGTGTGTGTGTGTGT 784
 DB 236 IleTrpLysAlaAlaLysPheArgPheGlyArgArgArgArgAlaValValProLeuPro 255
 QY 785 GCCACCATGCAAGTGAAGTCCAAAGTAAAGGAAGCACTGTAGTGAAGTGAAGTGAAGT 844
 DB 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
 QY 845 ACGGCACATGCAAGCAACGCTGCTTCAGGTGAGCGGGGACTCTCTGGCGGAGCGAG 904
 DB 272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
 QY 905 AAGGAGAGCGGAGCAGCATGATGTTGGGAATTCGATTGGCGTGTGTGTGTGTGTGTGT 964
 DB 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
 QY 965 ATCCCTCTTCTCCACCGAATCATACGCCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1024
 DB 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProIle 331
 QY 1025 TGGAAAAGCATATTTCTGTGGCTTGGCTACTCCAATTCTTTTCTTCAACCCCTGATTAC 1084
 DB 332 TrpLysSerIlePheLeuTrpLeuGlyTrpSerAsnSerPhePheAsnProLeuIleTrp 351
 QY 1085 ACAGCTTTTAAACAAGAACTACAACATGCTTCAAGAGCGCTCTTACTTAAGCAGAGA 1141
 DB 352 ThrAlaPheAsnLysAsnTrpAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
 RESULT 11
 ADC86157
 ID ADC86157 standard; protein; 606 AA.
 XX
 AC ADC86157;
 XX
 DT 01-JAN-2004 (first entry)
 XX

DE Human GPCR protein SEQ ID NO:610.
 KW human; GPCR; guanosine triphosphate-binding protein coupled receptor;
 KW gene therapy.
 OS Homo sapiens.
 XX EPI270724-A2.
 PN 02-JAN-2003.
 PD 18-JUN-2002; 2002EP-00013517.
 XX 18-JUN-2001; 2001JP-00246789.
 PF (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
 PR (ADSC-) CENT ADVANCED SCI & TECHNOLOGY INCUBATIO.
 XX Suwa M, Asai K, Akiyama Y, Aburatani H;
 XX WPI; 2003-315783/31.
 DR N-PSDB; ADC86156.
 XX
 PT New polynucleotide, useful for preparing a composition for treating a
 PT patient in need of increased or suppressed activity or expression of the
 PT guanosine triphosphate-binding protein coupled receptor.
 XX
 PS Claim 2; SEQ ID NO 610; 28pp; English.
 XX
 CC The invention relates to a novel polynucleotide encoding a guanosine
 CC triphosphate-binding protein coupled receptor (GPCR). A polynucleotide of
 CC the invention may have a use in gene therapy. The polynucleotide and
 CC polypeptide are useful for preparing a composition for treating a patient
 CC in need of increased or suppressed activity or expression of the
 CC guanosine triphosphate-binding protein coupled receptor. The protein
 CC sequences shown in ADC8549-ADC87617 represent GPCR's of the invention.
 XX
 SQ Sequence 606 AA;
 Alignment Scores:
 Pred. No.: 2,16e-92 Length: 606
 Score: 1296.50 Matches: 274
 Percent Similarity: 83.33% Conservative: 6
 Best Local Similarity: 81.55% Mismatches: 24
 Query Match: 60.19% Indels: 33
 DB: 7 Gaps: 5
 US-09-976-782-15 (1-1152) x ADC86157 (1-606)
 QY 3 CCATGGAGCGCGTAGCTTTTCAGTGGCCACCGCGCGGTTCCTTCCTCGGACCGC 62
 Db 254 ProTprArgProLeuAlaPheGlnTrpProProAlaLeuProLeuProTprPaspPro 273
 QY 63 AGACCAGCAGC-GGACCGGACCCCAAGCCGAGAGGATACCTCGGTTCCACCCGAGC 121
 Db 274 ArgProAlaAlaGlyProGlyThrProSerProArgGlyIleLeuGlySerThrProSer 293
 QY 122 GCGCGCTCTCTGCGGCGCGGAGGCGCCCTTCCTGCTTCACGGTCTCGTGGTACG 181
 Db 294 GlyAlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValThr 313
 QY 182 CTGCTAGTGTCTGATCGCTGCGCACTTCTCTGTGGAACTCGTGGTTCCGGTCCACCATC 241
 Db 314 LeuLeuValLeuLeuLeuAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIle 333
 QY 242 CCGCGGTCCTGCTTCACCGGTCGCGCATACCTTGTGGCTCGACGCGCTCTCG 301
 Db 334 ProArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSer 353
 QY 302 GACGAACCTAGTGGAGCGCTGGCATGCGACCGAGCTGGCGAGTACGCTTCACCGGG 361
 Db 354 AspGluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGly 373

QY 362 CGACGTGGCTGCTGGCGCGAGCGCTGTGCCACAGTGTGGATCTCTCTCGACCGGAGCC 421
 Db 374 ArgArgArgLeuLeuGlyArgSerLeuCysHisProValProArgValAspLeuLeuPro 393
 QY 422 TGTCTGTGCTCCCCCGCGCTCGGGAACGTGGCGGCATCGCCCTGGCGCGCACGGG 481
 Db 394 ArgLeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGly 413
 QY 482 GCCATCACAGCGCACTGCAGCACACGCTGGCCACCCCGAGCGCGCTCGTGTCTCATG 541
 Db 414 AlaIleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMet 433
 QY 542 ATCGCGTCCGCGCGGTGCGCTCGCGCTCATCGCCCTCGCGCGCTCTCTTGGCGGG 601
 Db 434 IleAlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArg 453
 QY 602 GCGAGGTGTGCGAGCTCGCTCCAGCGCTCCAGGTGAGCCCGGAAACCTCTCTATGCC 661
 Db 454 GlyGluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAla 473
 QY 662 GCCTTCTCCACCGCGCGCTTCCACCTGCGCTTGGCGTGGCTGCTGTCTTGTCTACCGG 721
 Db 474 AlaPheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArg 493
 QY 722 AAGATCTACGAGCGCGCCAGTTCGTTTCGCGCGCGCGCGAGAGCTGTGTCGCGCTG 781
 Db 494 LysIleTyrGluAlaAlaLysPheArgPheGlyArgArgAlaGluAlaValLeuProLeu 513
 QY 782 CCGGCCACCATGCGAGGTGAGGTCCAAAGGTA----- 811
 Db 514 ProAlaThrMetGlnValArgGlyGlyLeuArgAsnValAlaMetGlyLysArgLeuLeu 533
 QY 812 ---AAGGAAGCACCTGATGAGGCTG---AAGTGG-----TGTTCACGGCACATGCG 856
 Db 534 GlulysGluAlaAlaSer-ArgMetGlyGluTrpAlaGluAlaCysThrAsnGlyAlaAr 553
 QY 857 AAAGCAACGCGTGTCT-----TCCAGGTGAGCGGG 886
 Db 553 gAlaGlnArgSerProGlyAlaHisGluAspLysPheAlaIleSerSerGluAlaGln 573
 QY 887 GACTCTCTGGC-----GGAGCAGAGGAGAGGC 914
 Db 573 yThrGluGlyLeuValThrGlySerProGlyThrGlnValArgGly 588
 RESULT 12
 AAU79252
 ID AAU79252 standard; protein; 357 AA.
 XX AC AAU79252;
 XX DT 30-JUL-2002 (first entry)
 XX DE Human 5-hydroxytryptamine receptor 5A (HTR5A).
 XX KW Human; 5-hydroxytryptamine receptor 5A; HTR5A; serotonin;
 XX KW neuroprotective; neurological disease; depression; epilepsy;
 XX KW gene therapy; single nucleotide polymorphism; haplotype pair.
 XX OS Homo sapiens.
 XX PN WO200222887-A1.
 XX PD 21-MAR-2002.
 XX PF 17-SEP-2001; 2001WO-US029210.
 XX PR 15-SEP-2000; 2000US-0233051P.
 XX PA (GENA-) GENAISSANCE PHARM INC.
 XX PI Kazemi A, Koshy B, Sanchis A, Tirrell C;
 XX WPI; 2002-393978/42.

DR N-PSDB; ABK50434.

XX Novel genetic variants of 5-Hydroxytryptamine (Serotonin) Receptor 5A

PT isogenes, useful for improving efficiency and reliability in drug

PT development for treating neurological diseases.

XX

PS Claim 31; Fig 4; 134pp; English.

XX

CC The invention relates to single nucleotide polymorphisms in the gene

CC encoding human 5-hydroxytryptamine (serotonin) receptor 5A (HTR5A). A

CC method for haplotyping the HTR5A gene in an individual comprises

CC identifying the nucleotide at one or more polymorphic sites and

CC determining whether one of the copies of the gene is defined by one of

CC the HTR5A haplotypes given in the specification or whether both copies

CC are defined by a haplotype pair. This method is useful in genotyping,

CC whereby all possible haplotype pairs can be assigned to specific

CC genotypes. An association between a trait and a haplotype or haplotype

CC pair of the HTR5A gene can be identified by comparing the frequency of

CC the haplotype or haplotype pair in a population exhibiting the trait with

CC the frequency of the haplotype or haplotype pair in a reference

CC population, where a higher haplotype frequency in the trait population

CC indicates the trait is associated with the haplotype or haplotype pair.

CC HTR5A and its corresponding DNA are used for studying the expression and

CC function of HTR5A, and in screening for candidate drugs to treat diseases

CC related to HTR5A activity, such as neurological disorders, including

CC depression and epilepsy. This sequence represents the human HTR5A

CC polypeptide

XX

SQ Sequence 357 AA;

Alignment Scores:

Pred. No.:	2,73e-74	Length:	357
Score:	1063.50	Matches:	226
Percent Similarity:	72.98%	Conservative:	36
Best Local Similarity:	62.95%	Mismatches:	86
Query Match:	49.37%	Indels:	11
DB:	5	Gaps:	6

US-09-976-782-15 (1-1152) x RAU79252 (1-357)

QY 65 ACCAGCAGCGGACCGGGACCCCAAGCCCGGAGGAGTACTCGTTGCGACCCCGGCGGC 124

DB 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyLysAsp 27

QY 125 GCGGCTCTCGGCGGCGGCGGCGGCTTCTGCTCTTCCAGCTGCTGCTGCTGCTGCTG 184

DB 28 AspLeuArgProSer---SerProLeuLeuSerValPheGlyValLeuLeuLeuThrLeu 46

QY 185 CTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 244

DB 47 LeuGlyPheLeuValAlaAlaThrPheAlaTrpAsnLeuLeuValLeuAlaThrLeuLeu 66

QY 245 CGGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 304

DB 67 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86

QY 305 GAACTAGTGCAGCTGCGGAGTGCACCGAGCTGCGAGTGCAGTGCAGTGCAGTGCAGTGC 364

DB 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105

QY 365 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424

DB 106 ArgTrpGlnLeuGlyArgLeuCysGlnLeuTrpLeuAlaCysAspVal-----122

QY 425 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484

DB 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSer 142

QY 485 ATCAGACGCGACCTGCGAGCAGCTGCGACCGCGAGCGCGCTGCTGCTGCTGCTGCTGCTG 544

DB 143 IleThrArgHisMetGluTrpThrLeuArgThrArgLysCysValSerAsnValMetIle 162

QY 545 GCGCTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604

163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182

QY 605 GAGGTGTGCGAGCTCGGCTCCAGCGCTGCGAGGTGAGCGCGGAGACCTCTCTATGCGGCC 664

DB 183 GluThrTyrSerGluGlySerGluCysGlnValSerArgGluProSerTyrAlaVal 202

QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 724

DB 203 PheSerThrValGlyAlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLys 222

QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCGC---CGCGCGCGGAGAGCTGCTGCTGCTGCTG 781

DB 223 IleTyrLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242

QY 782 CGCGCCACCATCGAGTGGTCCCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGGTG 841

DB 243 SerGluAlaValGluValLysAspSerAlaLys-----GlnProGlnMetVal 258

QY 842 TTCACGCGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGAG 901

DB 259 PheThrValArg---HisAlaThrValThrPheGlnProGluGlyAspThrTrpArgGlu 277

QY 902 CAGAGGAGAGCGAGCGAGCCATGATGTGGGAATCTCTGATGGCGGTGTTGCTGCTGCTG 961

DB 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297

QY 962 TGGATCCCTCTCTCTGAGCGGAACCTCATCAGCCCACTCTGTCGCTGAGCGCTGCCCCC 1021

DB 298 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317

QY 1022 ATCTGGAAGACATATTTCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1081

DB 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIle 337

QY 1082 TACACAGCTTTTAAACAGAACTACAAATGCTTCAAGAGCTCTTTTACTAAGCAG 1138

DB 338 TyrThrAlaPheAsnLysAsnTyrAsnSerAlaPheLysAsnPhePheSerArgGln 356

RESULT 13

ABP81980

ID ABP81980 standard; protein; 357 AA.

XX

AC ABP81980;

XX

DT 04-MAR-2003 (first entry)

XX

DE Human 5-HT5A receptor protein SEQ ID NO:446.

XX

KW G protein-coupled receptor; GPCR; antigenic peptide; gene therapy;

KW G protein-coupled receptor; modulator; antibody; immune-related disease;

KW growth-related disease; cell regeneration-related disease; AIDS; cancer;

KW immunological-related cell proliferative disease; autoimmune disease;

KW Alzheimer's disease; atherosclerosis; infection; osteoarthritis; allergy;

KW osteoporosis; cardiomyopathy; inflammation; Crohn's disease; diabetes;

KW graft versus host disease; Parkinson's disease; multiple sclerosis; pain;

KW psoriasis; anxiety; depression; schizophrenia; dementia; memory loss;

KW mental retardation; epilepsy; asthma; tuberculosis; obesity; nausea;

KW hypertension; hypotension; renal disorder; rheumatoid arthritis; trauma;

KW ulcer.

OS Homo sapiens.

XX

XX WO200261087-A2.

XX

PD 08-AUG-2002.

XX

PF 19-DEC-2001; 2001WO-US050107.

XX

PR 19-DEC-2000; 2000US-0257144P.

XX

PA (LIFE-) LIFESPAN BIOSCIENCES INC.

XX

PR 18-JAN-2002; 2002US-0349511P.
 PR 28-FEB-2002; 2002US-0360500P.
 PR 15-MAR-2002; 2002US-0365041P.
 PR 19-APR-2002; 2002US-0374063P.
 PR 14-AUG-2002; 2002US-0403468P.
 PR 27-SEP-2002; 2002US-0414262P.
 PR 21-OCT-2002; 2002US-0419986P.
 PR 05-NOV-2002; 2002US-0423809P.
 PR 26-NOV-2002; 2002US-0429797P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Silos-Santiago I, Karicheti V;
 XX
 DR WPI: 2003-598705/56.
 DR N-PSDB; AAD58496.
 XX
 PT Identifying a compound for treating urological disorders, for example
 PT urinary incontinence by assaying the ability of the compound to modulate
 PT the nucleic acid expression or polypeptide activity.
 XX
 PS Disclosure; Page 187; Opp; English.
 XX
 CC The present relates to a method for identifying a compound for treating
 CC urological disorders e.g., urinary incontinence including overactive/
 CC oversensitive bladder, overflow urinary incontinence, stress urinary
 CC incontinence caused by dysfunction of the bladder, urethra or central or
 CC peripheral nervous system, prostatitis, benign prostatic hyperplasia,
 CC cancer of the prostate or kidney disorders. The method is also useful for
 CC modulating hyperplasia in a cell and treating a subject having a
 CC urological disorder. The invention is also used in gene therapy. The
 CC present sequence is human GPCR known as 5-hydroxytryptamine 5A (5-HT-5A)
 CC (serotonin receptor)
 XX
 SQ Sequence 357 AA;

Alignment Scores:
 Pred. No.: 2,73e-74 Length: 357
 Score: 1063.50 Matches: 226
 Percent Similarity: 72.98% Conservative: 36
 Best Local Similarity: 62.95% Mismatches: 86
 Query Match: 49.37% Indels: 11
 DB: 7 Gaps: 6

US-09-976-782-15 (1-1152) x AA#38595 (1-357)

QY 65 ACACAGCGGACCCCGGACCCCAAGCCGAGGAGTACTCGTTGACCCCGAGCGGC 124
 DB 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyIysAsp 27
 QY 125 GCGGTCTCTCGCGGCGAGGCGGCCCTTCTCTGTCTTACAGCTCTGTGGTGAGCGTG 184
 DB 28 AspLeuArgProSer---SerProLeuSerValPheGlyValLeuLeuThrLeu 46
 QY 185 CTAGTGCTCTGCTGCTCCACTTTCCTGTGGAACCTGCTGCTCCGCTCCACATCCG 244
 DB 47 LeuGlyPheLeuValAlaAlaThrPheAlaTrpAsnLeuValLeuAlaThrLeu 66
 QY 245 CGGTCCTGCTCTCCACCGGTCGCGATACCTTGGTGGCTCGAGCGGCTCTCGAC 304
 DB 67 ArgValArgThrPheHisArgValProHisAsnLeuValAlaSerMetAlaValSerAsp 86
 QY 305 GAACTAGTGGCGGCTGGCGATGCCACCGAGCTGGCGAGTGGCTGCGACCGGCGCA 364
 DB 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 QY 365 CGTGGCTCTGGCGGAGCCTGTGCCACGTGTGGATCTCTTCGACCGCGAGCCTGT 424
 DB 106 ArgTrpGlnLeuGlyArgArgLeuGlyGlnLeuTrpIleAlaCysAspVal----- 122
 QY 425 CTGTGCTGCGCGGCTCGGAGACGTGGCGCATCGCCCTGGCGCGGCGGCGGC 484
 DB 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTyrTrpSer 142

QY 485 ATCACACGGCACCTGCAGCACACACGCTGCGCACCGCGCGCGCTCGTTCATGATC 544
 DB 143 IleThrArgHisMetGluTyrThrLeuArgThrArgLysCysValSerAsnValMetIle 162
 QY 545 GCGCTGCGCCGGTGCCTGCGCGCTCATCGCCCTCGCCCGCTGTCTTTGCCCGGGC 604
 DB 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 QY 605 GAGGTGTGCACGCTCGCGCTCCAGCGCTGCCAGGTGAGCGGAGACCTCTATGCGGC 664
 DB 183 GluThrTyrSerGluGlySerGluGluCysGlnValSerArgGluProSerTyrAlaVal 202
 QY 665 TTCTCCACCGCGCGCTTCACCTGCGCTTGGCGTGTGGCGTGTCTCTACCGGAAG 724
 DB 203 PheSerThrValGlyAlaPheTyrLeuProLeuCysValValLeuPheValTyrTrpLys 222
 QY 725 ATCTACGAGCGCGCAAGTTTCGTTTCGCG---CGCCGCGGAGAGCTGTGTCGCTTG 781
 DB 223 IleTyrLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProIle 242
 QY 782 CCGGCCACCATGCAGGTGAGTCCCAAGGTAAAGGAAGCACCTGATGAGGCTGAAGTGTG 841
 DB 243 SerGluAlaValGluValLysAspSerAlaLys-----GlnProGlnMetVal 258
 QY 842 TTTCAGCGCACATTGCAAGCAACGGTGTCTCTCCAGGTGAGCGGCGACTCTCGCGGAG 901
 DB 259 PheThrValArg---HisAlaThrValThrPheGlnProGluGlyAspThrTrpArgLys 277
 QY 902 CAGAGGAGAGCGGACGACCCATGATGTTGGGAAATCTGATGCGCTGTGTTGCTGTGC 961
 DB 278 GlnLysGluGlnArgAlaAlaLeuMetValGlyIleLeuIleGlyValPheValLeuCys 297
 QY 962 TGGATCCCTCTCTCTGACGGAACTCATCAGCCCACTGTGCTGCTGAGCTGCCCCC 1021
 DB 298 TrpIleProPhePheLeuThrGluLeuIleSerProLeuCysSerCysAspIleProAla 317
 QY 1022 ATCTGMAAAGCATATTTCTGTGGCTTGGCTACTTCCAAATCTTCTTCAACCCCTGATT 1081
 DB 318 IleTrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIle 337
 QY 1082 TACACAGCTTTTAAACAAGAACTACAACTGCTTCAAGAGCCTCTTTACTAAGCAG 1138
 DB 338 TyrThrAlaPheAsnLysAsnTyrAsnSerAlaPheLysAsnPhePheSerArgGln 356
 RESULT 15
 ADE56996
 ID ADE56996 standard; protein; 357 AA.
 XX
 AC ADE56996;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Human Protein NP_076917, SEQ ID NO 2851.
 XX
 KW Human; pain; neuronal tissue; gene therapy;
 KW spinal segmental nerve injury; chronic constriction injury; CCI;
 KW spared nerve injury; SNI; Chung.
 XX
 OS Homo sapiens.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-033347P.
 XX
 XX (GEHO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:06:29 ; Search time 158 Seconds
(without alignments)
4600.974 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgccatgagccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 2034082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlip
-Q=/cgcn2/1/USPTO.spool_p/US09976782/runat_31082004_115951_10315/app_query.fasta_1.1351
-DB=SPTEMBL_25 -QFMT=faстан -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09976782 -CGN_1_213 -runat_31082004_115951_10315 -NCFU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPTEMBL_25.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1050.5	48.8	346	13 Q7Z32	Q7Z32 brachydanio

2	615	28.6	247	4 Q86UN1	Q86un1 homo sapien
3	451	20.9	390	6 Q3N2B7	Q3n2b7 gorilla gor
4	438	20.3	386	11 Q99MB0	Q99mb0 mesocricetu
5	437.5	20.3	423	6 Q8HYB0	Q8hyb0 canis famli
6	434.5	20.2	464	5 Q9GQ54	Q9gq54 aedes aegyp
7	433.5	20.1	492	5 Q76267	Q76267 aplysia cal
8	432	20.1	422	6 Q9N297	Q9n297 gorilla gor
9	431.5	20.0	445	5 Q18512	Q18512 caenorhabdi
10	431	20.0	421	11 Q8BZP1	Q8bzp1 mus musculu
11	429	19.9	422	6 Q9N296	Q9n296 pongo pygma
12	428	19.9	422	6 Q9N298	Q9n298 pan troglod
13	426.5	19.8	390	6 Q9GLP5	Q9glp5 sus scrofa
14	423.5	19.7	374	11 Q8BUW7	Q8buw7 mus musculu
15	423	19.6	421	11 Q8BGS4	Q8bgs4 mus musculu
16	421.5	19.6	408	13 Q98998	Q98998 xenopus lae
17	419.5	19.5	377	6 Q9N263	Q9n263 sus scrofa
18	419	19.5	447	6 Q8SPH2	Q8sph2 sus scrofa
19	418	19.4	567	5 Q8MX83	Q8mx83 aplysia cal
20	413.5	19.2	365	6 Q9N2D8	Q9n2d8 gorilla gor
21	412.5	19.2	365	4 Q9P204	Q9p204 homo sapien
22	411.5	19.1	365	6 Q9N2D9	Q9n2d9 pan troglod
23	409.5	19.0	365	6 Q9N2D7	Q9n2d7 pongo pygma
24	408	18.9	363	6 Q9N2B6	Q9n2b6 pan troglod
25	408	18.9	363	6 Q9N2B5	Q9n2b5 gorilla gor
26	408	18.9	470	11 P97842	P97842 rattus norv
27	405	18.8	363	6 Q9N2B4	Q9n2b4 pongo pygma
28	405	18.8	366	6 Q9GKI2	Q9gki2 sus scrofa
29	401	18.6	439	5 Q86LL7	Q86ll7 haemochus
30	399	18.5	417	13 Q8JG03	Q8jg03 fugu rubrip
31	397.5	18.5	399	5 Q9NG02	Q9ng02 apis mellif
32	397	18.4	419	5 Q77254	Q77254 boophilus m
33	394	18.3	477	5 Q9BMA9	Q9bma9 mamestra br
34	390	18.1	415	13 Q8JG69	Q8jg69 brachydanio
35	389	18.1	388	13 Q90WF4	Q90wf4 brachydanio
36	388.5	18.0	437	13 Q7TIA2	Q7tia2 brachydanio
37	382	17.7	414	6 Q9GKA0	Q9gka0 canis famli
38	382	17.7	414	6 Q9GK99	Q9gk99 canis famli
39	380	17.6	518	6 Q9MYI8	Q9myi8 oryctolagus
40	378	17.5	459	11 Q8CGI5	Q8cgi5 mus musculu
41	378	17.5	515	11 Q9DBL0	Q9db10 mus musculu
42	373	17.3	465	4 Q86TH8	Q86th8 homo sapien
43	372.5	17.3	402	5 Q44198	Q44198 apis mellif
44	371.5	17.2	454	13 Q8AWE1	Q8awe1 brachydanio
45	370.5	17.2	448	11 Q925K7	Q925k7 mus musculu

ALIGNMENTS

RESULT 1

Q7Z32 PRELIMINARY; PRT; 346 AA.
ID Q7Z32
AC Q7Z32
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SI:ZC12P8.3 (Novel protein similar to human 5-hydroxytryptamine
receptor 5A (HTR5A)).
GN SI:ZC12P8.3.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
[1]
RN SEQUENCE FROM N.A.
RP Corby N.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
EMBL: AL772146; CAD61100.1; -
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.

Score: 615.00 Matches: 143
 Percent Similarity: 67.35% Conservativity: 22
 Best Local Similarity: 58.37% Mismatches: 74
 Query Match: 28.55% Indels: 6
 DB: 4 Gaps: 4

US-09-976-782-15 (1-1152) x Q86UN1 (1-247)

QY 65 ACCAGCAGCGGACCGCGGACCCAGCCAGCGAGGAGGATCTCGTTGACCCGAGCGGC 124
 DB |||||
 QY 8 ThrSerPheSerLeuSerThrProSerProLeuGluThrAsnHisSerLeuGlyLysAsp 27
 DB |||||
 QY 125 GCCGTCCTGCGCGGCGGAGGCGCCCTCTCTCTTCACGCTGCTGCTGCTGCTGCTG 184
 DB |||||
 QY 28 AspLeuArgProSer---SerProLeuLeuSerValPheGlyValLeuLeuLeuLeuLeu 46
 DB |||||
 QY 185 CTAGTGCTGCTGATCGCTCCACCTCTCTCTGGAACCTGCTGCTGCTGCTGCTGCTGCTG 244
 DB |||||
 QY 47 LeuGlyPheLeuValAlaAlaThrPheAlaTrpAsnLeuLeuValLeuAlaThrIleLeu 66
 DB |||||
 QY 245 CGGCTCCGCTGCTTCCACCGCTGCGCATAACTTGGTGGCTGCGACGCGCTCTCGGAC 304
 DB |||||
 QY 67 ArgValArgThrPheHisArgValProHisnLeuValAlaSerMetAlaValSerAsp 86
 DB |||||
 QY 305 GAATAGTGGCAGCTGCGGATGATCCACCGACCTGCGGAGTGAAGTCTGTCGACCGGCGGA 364
 DB |||||
 QY 87 ValLeuValAlaAlaLeuValMetProLeuSerLeuValHisGluLeuSer---GlyArg 105
 DB |||||
 QY 365 CGTGGCTGCTGGCGGAGCTGTCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 424
 DB |||||
 QY 106 ArgTrpGlnLeuGlyArgLeuGlyCysGlnLeuTrpIleAlaCysAspVal----- 122
 DB |||||
 QY 425 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 484
 DB |||||
 QY 123 LeuCysCysThrAlaSerIleTrpAsnValThrAlaIleAlaLeuAspArgTrpSer 142
 DB |||||
 QY 485 ATCACAGGACCTGCGACACACGCTGCGACCGCGCGCTGCTGCTGCTGCTGCTGCTGCTG 544
 DB |||||
 QY 143 IleThrArgHisMetGluTrpThrLeuArgThrArgLysCysValSerAsnValMetile 162
 DB |||||
 QY 545 GCGTCCCGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
 DB |||||
 QY 163 AlaLeuThrTrpAlaLeuSerAlaValIleSerLeuAlaProLeuLeuPheGlyTrpGly 182
 DB |||||
 QY 605 GAGTGTGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664
 DB |||||
 QY 183 GluThrTrpSerGluGlySerGluGluCysGlnValSerArgGluProSerTrpAlaVal 202
 DB |||||
 QY 665 TTCTCCACCGCGGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724
 DB |||||
 QY 203 PheSerThrValGlyAlaPheTrpLeuProLeuGlyValLeuPheValTrpLys 222
 DB |||||
 QY 725 ATCTACGAGCGGCGCAAGTTTCTGTTTCGGC---CGCGCGGAGAGTGTGCTGCTGCTGCTG 781
 DB |||||
 QY 223 IleTrpLysAlaAlaLysPheArgValGlySerArgLysThrAsnSerValSerProile 242
 DB |||||
 QY 782 CCGGCCACCATGCGAG 796
 DB |||||
 QY 243 SerGluAlaValGlu 247
 DB |||||

RESULT 3

Q9N2B7 PRELIMINARY; PRT; 390 AA.
 ID Q9N2B7
 AC Q9N2B7;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE 5-hydroxytryptamine (Serotonin) receptor 1B.
 GN HTR1B.
 OS Gorilla gorilla (gorilla).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
 OC NCBI_TaxID=9593;

Alignment Scores:
 Pred. No.: 4.45e-22 Length: 390
 Score: 451.00 Matches: 127
 Percent Similarity: 48.35% Conservativity: 63
 Best Local Similarity: 32.32% Mismatches: 147
 Query Match: 20.94% Indels: 56
 DB: 6 Gaps: 13

US-09-976-782-15 (1-1152) x Q9N2B7 (1-390)

QY 59 CCCGAGACCCAGCGGACCGGACCCAGCCAGCGGAGGATCTCGTTGACCCCGG 118
 DB |||||
 QY 10 ProPro***ProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerSerAlaPro 29
 DB |||||
 QY 119 AGC-----GGCGCGCTCTCCCGGCGGA 142
 DB |||||
 QY 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleAlaLeuPro----- 47
 DB |||||
 QY 143 GGGCGCGCTTCTCTGCTTTCACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 202
 DB |||||
 QY 48 -----TprLysValLeuLeuValMetLeuLeuAlaLeuIleThrLeu 61
 DB |||||
 QY 203 GCACCTTTCCTGTTGGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 262
 DB |||||
 QY 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
 DB |||||
 QY 263 CGCGTGGCGCATAACTTGGTGGCTGCTGCGACCGCGCTCTCGGACGAACCTAGTGGAGCGGTG 322
 DB |||||
 QY 82 ThrProAlaAsnTyrLeuIleAlaSerLeuAlaValThrAspLeuLeuValSerIleLeu 101
 DB |||||
 QY 323 CGGATGCCACCGGCTGCGGAGTAGAGTGTGTCGACCGGCGGACGCTGCTGCTGCTGCTGCTG 382
 DB |||||
 QY 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgTrpThrLeuGlyGln 119
 DB |||||
 QY 383 AGCGTGTGCCACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 442
 DB |||||
 QY 120 ValValCysAspPheTrpLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
 DB |||||
 QY 443 CTCGGGAACCTGCGCGCATCTGCGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 502
 DB |||||
 QY 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
 DB |||||
 QY 503 CACAGCTGCGCACCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 562
 DB |||||
 QY 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetileAlaLeuValTrpValPhe 176
 DB |||||
 QY 563 TCGGCGCTCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 622
 DB |||||
 QY 177 SerIleSerIleSerLeuProPro---PhePheTrpArgGlnAlaLysAlaGluGluGlu 195
 DB |||||
 QY 623 CTCAGCGCTGCGAGGTGACCGGGAACCC---TCCTATCGCGCTTCTCCACCGCGGCGG 679
 DB |||||

[1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-gorilla-UI;
 RA Kitano T., Kobayakawa H., Saitou N.;
 RT "Silver Project";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

EMBL; AB041372; BAA94457.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; F:receptor activity; IEA.
 DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPTOR FL2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 390 AA; 43565 MW; 37D54B95E3277FC3 CRC64;

OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes.
 OX NCBI_TaxID=71159;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Pietrantonio P.V., Jagger C.L.;
 RT "Cloning of a serotonin receptor 5HT7-like cDNA from mosquito Aedes
 RT aegypti (Diptera: Culicidae) excretory and respiratory systems.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF296125; AAG49292.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; P:receptor activity; IEA.
 DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G-PROTEIN RECEPT_F1_1; 1.
 DR PROSITE; PS0262; G-PROTEIN RECEPT_F1_2; 1.
 KW G-protein coupled receptor; Receptor; Transmembrane.
 SQ SEQUENCE 464 AA; 50821 MW; 3F432DA22E0345E4 CRC64;
 Alignment Scores:
 Pred. No.: 5 69e-21 Length: 464
 Score: 434.50 Matches: 118
 Percent Similarity: 49.35% Conservative: 71
 Best Local Similarity: 30.81% Mismatches: 163
 Query Match: 20.17% Indels: 31
 DB: 5 Gaps: 9
 US-09-976-782-15 (1-1152) x Q9GQ54 (1-464)
 QY 11 GCCGCTAGCTTTAGTGCCACCGCGCGTGTGCCTTCCCTTGCCTGGGA----- 58
 Db 58 SerSerSerAlaSerValSerAlaThrAlaIleAlaThrSerLeuProAlaLeuValAsp 77
 QY 59 -----CCGAGACACGAGCGGACCGCGGACCCCAAGCCGAGGAGTACTCGGTCG 112
 Db 78 ArgLeuThrProThrSerSer-----ThrSerSerLeuLeuAspGluLeuGlyGly 94
 QY 113 ACCCGAGCGCGCGTCTCGCGCGCGCGCGCTTCTCTGTCTTACGGFCCTG 172
 Db 95 ThrSerGluSerSer-----ProAlaGluProValAsnValLeuThrIleGln 110
 QY 173 GTGGTACGCTGTAGTGTGCTGATCGCT-----GCCACTTCTGTGGAACTCG 223
 Db 111 ThrIleValIleSerIleValLeuLeuAlaValIleIleGlyThrIleValGlyAsnVal 130
 QY 224 CTGTTCCGCTACCATCCGCGGTCGCGCTTCCACCGCGTCCGCGATACCTAGTGTG 283
 Db 131 LeuValCysValAlaValCysLeuValArgLysLeuArgArgProCysAsnTyrLeuLeu 150
 QY 284 GCTCGACGCGCTCGGACGAACTAGTGGAGCGCTGCGATGCGACCGCGCGCGGCG 343
 Db 151 ValSerLeuAlaIleSerAspLeuCysValAlaValLeuValMetProProAlaLeuLeu 170
 QY 344 AGTGAAGCTCGACCGCGGACGCTCGGCTGCTGGCGGAGGCGCTGTGCCACGCTGTGATC 403
 Db 171 TyrGluValLeuGluGluTrpLys-----PheGlyThrValPheCysAspIleTrpVal 188
 QY 404 TCTTCGACCGCGGACCTGTCTGTCTGCTCCCGCGCGCTCGGAACTGGCGGCGCATC 463
 Db 189 SerPheAspVal-----LeuSerCysThrAlaSerIleLeuAsnLeuCysAlaIle 205
 QY 464 GCCTGGCGCGCGGCGGCGCATCACACGCACTGCAGCACACGCTGGCGCCACCGCAGC 523
 Db 206 SerValAspArgTyrTrpAlaIleThrLysProLeuGluTyrGlyValLysArgThrPro 225
 QY 524 CCGCGCTGCTGTCTATGATCGCGCTGCGCGCGGCTGCGCGCTGTCTATCGCGCTCGG 583
 Db 524 CCGCGCTGCTGTCTATGATCGCGCTGCGCGCGGCTGCGCGCTGTCTATCGCGCTCGG 583

Db 226 ArgArgMetIleAlaCysIleValLeuValTrpLeuValAlaAlaCysIleSerLeuPro 245
 QY 584 CGCGTGTCTTTGGCGGGGAGGTGTGCGACGCTCGCGTCCAGCGCTGCCAGGTGAGC 643
 Db 246 ProLeuLeuIleLeuGlyAsnGluHisMetThrAsnGlyGlnProSerCysSerValCys 265
 QY 644 CGGGAACCTCTCTATGCGCGCTTCTCCACCGCGCGCTTCCACCTGCGCTGCGGTG 703
 Db 266 GlnAsnPhePheTyrGlnIleTyrAlaThrLeuCysAlaPheTyrIleProLeuAlaVal 285
 QY 704 GTGCGCTTTGTACCGAAGATCTACAGGCGGCGCAAGTTTCGTTTCGCGCGCGCGG 763
 Db 286 MetLeuPheValTyrPheGlnIlePheArgAlaAlaArgArgIleValAsnGluGluLys 305
 QY 764 AGAGCTGTGCTGCGCTTGGCGGCCACCATGCGAGGTGAGGTCCAGGTAAAGGAGCACCT 823
 Db 306 ArgAlaGlnLysHisLeuGluThrAlaIleAsnGlySerAlaThrThrProGluLysLys 325
 QY 824 GATGAGGCT-----GAAGTGTGTTCACGGCACCATGTCAGAACACCGTGTCTTCCAG 877
 Db 326 LeuSerAlaGlyGlyThrValLeuValAlaThrProGlnHisLysArgLeuArgPheGln 345
 QY 878 GTGACGGGGACTCTCTGGCGGAGCAGAGGAGCGGAGCAGCAGCATGATGTTGGGAATT 937
 Db 346 LeuAla-----LysGluArgLysAlaSerThrThrLeuGlyIle 358
 QY 938 CTGATTGCGGTGTTGTGCTGTGCTGATGCTGCTTCTCTGACGGAACATCATGCCCA 997
 Db 359 IleMetSerAlaPheThrValCysTrpLeuProPheIleLeuAlaLeuValArgPro 378
 QY 998 CTCTGTGCTGCGAGCTGCGCCCTGAGAAAGCATATTTCTGTGGTGTGCTACTCC 1057
 Db 379 PheLeuGly---GluAspHisHisLeuLeuSerSerLeuPheLeuTrpLeuGlyTyrAla 397
 QY 1058 AATTCCTTCTCAAGCCCTGATTTACACAGCTTTTACAGAACTACACATGCTCTTC 1117
 Db 398 AsnSerLeuLeuAsnProIleIleTyrAlaThrLeuAsnArgAspPheArgLysProPhe 417
 QY 1118 AAGAGCCTC 1126
 Db 418 GlnGluIle 420
 RESULT 7
 O76267 PRELIMINARY; PRT; 492 AA.
 AC O76267;
 DT 01-NOV-1998 (TReMBLrel. 08, Created)
 DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
 DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)
 DE G-protein-coupled 5-hydroxytryptamine receptor.
 GN 5-HTA1.
 OS Aplysia californica (California sea hare).
 OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
 OC Apogastropoda; Heterobranchia; Euthyneura; Opisthobranchia; Anaspidia;
 OC Aplysioidae; Aplysiidae; Aplysia.
 OX NCBI_TaxID=6500;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98337861; PubMed=9671650;
 RA Angers A., Storzuk M.V., Duchaine T., Castellucci V.F.,
 RA Desgossiers L.,
 RT "Cloning and functional adenylation of an aplysia 5-HT receptor
 RT negatively coupled to adenylylate cyclase.";
 RL J. Neurosci. 18:5586-5593(1998).
 CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 DR EMBL; AF041039; AAC28786.1; -;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0004872; P:receptor activity; IEA.
 DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.

US-09-976-782-15 (1-1152) x Q9N297 (1-422)

```
QY 134 CGGGCCGAGGG-----CGCCCTTCTCT----- 157
Db 6 ProGlyGlnGlyAsnAsnThrThrSerProAlaProPheGluThrGlyGlyAsnThr 25
QY 158 -----GTC TTCACGGTCCTGGTGTCGACCGTGTAGTGTG----- 193
Db 26 ThrGlyIleSerAspValThrPheSerTyrGlnValIleThrSerLeuLeuLeuGlyThr 45
QY 194 CTGATCGTCGCACCTTCTCTGGAACCTGCTGCTCCGGTCCACATCCCGGGTCCGT 253
Db 46 LeuIlePheCysAlaValLeuGlyAsnAlaCysValValAlaAlaIleAlaLeuGluArg 65
QY 254 GCCTTCCACCGCTCGCCGATACCTTGGTGGCTCGACGGCGCTCTCGGACGAACCTAGTG 313
Db 66 SerLeuGlnAsnValAlaAsnTyrLeuIleGlySerLeuAlaValThrAspLeuMetVal 85
QY 314 GCAGCGCTCGGATGCCACCGACCTGGCGAGTGTGTCACCGGGGACGTCGGGTG 373
Db 86 SerValLeuValLeuPro-----MetAlaAlaLeuTyrGlnValLeuAsnLysTrpThr 103
QY 374 CTGGGCGGAGCTGTGTCACGTGTGATCTCTTCGACGCCGAGCCTGTCTGTCTGC 433
Db 104 LeuGlyGlnValThrCysAspLeuPheIleAlaLeuAspVal-----LeuCysCys 120
QY 434 CGCGCGGCTCGGGAACGTGGCGGCATCGCCCTGGCGGCGGCGGCATCACACGG 493
Db 121 ThrSerSerIleLeuHisLeuCysAlaIleAlaLeuAspArgTyrTrpAlaIleThrAsp 140
QY 494 CACCTCGACACACGCTGGCACCCCGACCGCGCTCTGTGTCATGATCGCGCTCGCC 553
Db 141 ProIleAspTyrValAsnLysArgThrProArgAlaAlaAlaLeuIleSerLeuThr 160
QY 554 CGGGTCCGCTCGCGCTCTCGCCCTCGCGCTCTCTTGGCGGGCGGAGTGTGC 613
Db 161 TrpLeuIleGlyPheLeuIleSerIleProProMetLeuGlyTrpArgThrProGluAsp 180
QY 614 GACGCTCGGCTCGACGCTGCCAGTGCAGCGGGAACCTCTATCGCGCTTCTCCAC 673
Db 181 ArgSerAspProAspAlaCysThrIleSerLysAspHisGlyTyrThrIleTyrSerThr 200
QY 674 CGGGCGCTTCCACCTCGCTGGTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 733
Db 201 PheGlyAlaPheTyrIleProLeuLeuLeuMetLeuValLeuTyrGlyArgIlePheArg 220
QY 734 GCGGCCAAGTTGGTTTCGCGCGC-----CGC 760
Db 221 AlaAlaArgPheArgIleArgLysThrValLysValGluLysThrGlyAlaAspThr 240
QY 761 CGGAGAGCTGTGCTGCGCTGTCGCGGC----- 787
Db 241 ArgHisGlyAlaSerProAlaProGlnProLysLysSerValAsnGlyGluSerGlySer 260
QY 788 -----ACATGCAAGTGGAGTCCAAG----- 808
Db 261 ArgAsnTrpArgLeuGlyValGluSerLysAlaGlyGlyAlaLeuCysAlaAsnGlyAla 280
QY 809 GTAAAGGAAGCACTCATGAGGCTGAAGTG--GTGTTCCAGCGCACAT----- 853
Db 281 ValArgGlnGlyAspAspGlyAlaAlaLeuGluValIleGluValHisArgValGlyAsn 300
QY 854 -----TGCAAAGCAACGGTG 868
Db 301 SerLysGluHisLeuProLeuProSerGluAlaGlyProThrProCys---AlaProAla 319
QY 869 TCTTCCAGTGAGCGGGAACCTCTCGCGGAGCAGAAG----- 907
Db 320 SerPheGluArgLysAsnGluArgAsnAlaGluAlaLysArgLysMetAlaLeuAlaArg 339
QY 908 GAGAGCGGAGCAGCCATGATGGTGGGAATTCATTTGGCGTGTTCGTGCTGCTGGATC 967
Db 340 GluArgLysThrValLysThrLeuGlyIleMetGlyThrPheIleLeuCysTrpLeu 359
```

```
QY 968 CCCTTCTCTCTGACGGAACCTCATCAGCCCACTCTGT-----GCCTGCAAGCTGCCCCC 1021
Db 360 ProPhePheIleValAlaLeuValLeuProPheCysGluSerSerCysHisMetProThr 379
QY 1022 ATCTGGAAGAAGCATATTTCTGTGGCTGTGGCTACTCAATCTTTCTTCAACCCCTGATT 1081
Db 380 LeuLeuGlyAlaIleIleAsnTrpLeuGlyTyrSerAsnSerLeuAsnProValIle 399
QY 1082 TACACAGCTTTTAAACAAGAACTACACAATGCCTTCAAGAGCTC 1126
Db 400 TyrAlaTyrPheAsnLysAspPheGlnAsnAlaPheLysLysIle 414
RESULT 9
O18512 PRELIMINARY; PRT; 445 AA.
AC O18512; Q9BKW9;
DT 01-JAN-1998 (trEMBLrel. 05, Created)
DT 01-JAN-1998 (trEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (trEMBLrel. 25, Last annotation update)
DE Serotonin receptor (Hypothetical protein).
GN Y22D7AR.13.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Khabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Olde B., McCombie R.W.;
RT Molecular cloning and functional expression of a serotonin receptor
RT from Caenorhabditis elegans.;
RL J. Mol. Neurosci. 0:0-0(1997).
RN [2]
RP SEQUENCE FROM N.A.
RA Dubnick M.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Madsen C.;
RT "The sequence of C. elegans cosmid Y22D7AR.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR EMBL; U15167; AAB66360.1; -.
DR EMBL; AC084154; AAK29875.2; -.
DR FIR; T42203; T42203.
DR WormPep; Y22D7AR.13; CE29370.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G-PROTEIN RECEPTOR FL1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEPTOR FL2; 1.
KW Hypothetical protein; G-protein coupled receptor; Receptor;
KW Transmembrane.
```



```
QY 833 GAAGTGTGTTCACGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCC 892
Db |||:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
283 SerProValTyrValAsnGlnValLysValArgValSer-----AspAla 297
QY 893 TGGCGGAGCAGCAAG-----GAGAGCGCAGCAGCAGCCATGATGGTCGGA 934
Db |||:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
298 LeuLeuGluLysLysLysLeuMetAlaAlaArgGluArgLysAlaThrLysThrLeuGly 317
QY 935 ATTCTGATGGCGTGTGTGCTGTGTGATGCCCTTCTCTCCAGCGGAACTCATCAGC 994
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
318 IleIleLeuGlyAlaPheIleValCysTrpLeuProPhePheIleIleSerLeuAlaMet 337
QY 995 CCATCTGT-----GCTGCAGCGCTGCCCCCATCTCGAAAGCATATTCTGTGGCTT 1048
Db |||:||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
338 ProlCysLysAspAlaCysTrpPheHisLeuAlaIlePheAspPheThrTrpLeu 357
QY 1049 GGTACTCCAAATCTTTCTTCAACCCCTGATTTACACAGCTTTTAAAGAACTACAAC 1108
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
358 GlyTyrLeuAsnSerLeuIleAsnProIleIleTyrThrMetPheAsnGluAspPheLys 377
QY 1109 AATGCCTTCAGAGCCCTC 1126
Db |||:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
378 GlnAlaPheHisLysLeu 383

RESULT 14
Q8BUW7 PRELIMINARY; PRT; 374 AA.
ID Q8BUW7 AC Q8BUW7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 5-hydroxytryptamine 1D receptor.
GN HTR1D.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Head;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK082016; BAC38393.1; -.
DR MGD; MGI:96276; Htr1d.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHOLOPSN.
DR PROSITE; PS00237; G PROTEIN RECEPT FL1; 1.
DR PROSITE; PS0262; G-PROTEIN RECEPT FL2; 1.
SQ SEQUENCE 374 AA; 41593 MW; 2FE947DFDBEF14D CRC64;
```

Alignment Scores:

Pred. No.:	3,05e-20	Length:	374
Score:	423.50	Matches:	119
Percent Similarity:	47.81%	Conservative:	67
Best Local Similarity:	30.59%	Mismatches:	150
Query Match:	19.66%	Indels:	53
DB:	11	Gaps:	12

US-09-976-782-15 (1-1152) x Q8BUW7 (1-374)

```
QY 53 CTGGAGCCGACAGCAGCAGCGGAGCCCGAGCCCAAGCCCGAGGGATACCTCGGTTCG 112
Db |||:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 MetSerProAsnGlnSerLeuGluGlyLeuProGlnGluAlaSerAsnArgSerLeu 20
QY 113 ACCCGAGCGCGCGTCTCTGCGGCGCAGCGGCCCTTCTGTCTTACAGGTCCTG 172
```

```
Db |||:||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
21 AsnValThrGlyAlaTrpAspPro-----GluValLeuGlnAlaLeu 34
QY 173 GTGGTGAAGTGTAGTGTCTGCTG-----ATCGTGCACATTTCTCTGTGGAACCTG 223
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
35 ArgIleSerLeuValValLeuSerValIleThrLeuAlaThrValLeuSerAsnAla 54
QY 224 CTGTTCCGGTCCACATCCCGCGGTGCGCTTCCACCGCGTCCGCGCAATAACTGGTG 283
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
55 PheValLeuThrIleLeuLeuThrLysLysLeuHisThrProAlaAsnTyrLeuIle 74
QY 284 GCCTCGACGGCGCTCTCGGACGAATACTAGTGCACGCTGGCGATGCAACCGAGCTGGCG 343
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
75 GlySerLeuAlaThrThrAspLeuValSerIleLeuValMetProIleSerIleAla 94
QY 344 AGTGAGTGTGCGACCGCGGCGACGTGCGTGTGGCGGAGCGCTGCCAGTGTGGATC 403
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
95 -----TyrThrThrThrArgTrpAsnPheGlyGlnIleLeuCysAspIleTrpVal 112
QY 404 TCCTTCGACGCGGAGCGCTGTCTGTGCTGCCCGCGCTCGGGAACAGTGGCGGCCATC 463
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
113 SerSerAsp-----IleThrCysThrAlaSerIleLeuHisLeuCysValIle 129
QY 464 GCCCTGGCGCGACGGGCGCATCACACGCGCATCTGCAGCACAGCTGCGCAGCCGCGAGC 523
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
130 AlaLeuAspArgTyrTrpAlaIleThrAspAlaLeuGluTyrSerLysArgArgThrAla 149
QY 524 CGCGCTGTGTTCATGATCGCTCGCGCGGTGCGCGCTCGGCGCTCATCCCTCGCG 583
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
150 GlyHisAlaAlaMetIleAlaAlaValTrpIleIleSerIleCysIleSerIlePro 169
QY 584 CCGCTGTCTTTGGCGGCGAGGTGTGCGACGCTCGGCTCCAGCGTGCAGGTG--- 640
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
170 Pro---LeuPheTrpArgGlnAlaThrAlaHisGluGluMetSerAspCysLeuValAsn 188
QY 641 AGCGGGAACCCCTCTATGCGCTTTCACCGCGCGCTTCCACCTGCGCTGCGCTGGC 700
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
189 ThrSerGlnIleSerTyrThrIleTyrSerThrCysGlyAlaPheTyrIleProSerIle 208
QY 701 GTGGTGGCTTGTCTACCGGAAGATCTACGAGCGCGCAAGTTTCGT----- 748
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
209 LeuLeuIleIleLeuTyrGlyArgIleTyrValAlaAlaArgSerArgIleLeuAsnPro 228
QY 749 -----TTCGCGCGCGCGAGAGCTGTGTGCGCTG----- 781
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
229 ProSerLeuTyrGlyLysArgPheThrThrAlaGlnLeuIleThrGlySerAlaGlySer 248
QY 782 -----CCGCGCACCATGCGAGGTGAGGTCCAAAGGTAAAGGAAGCACCT 823
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
249 SerLeuCysSerLeuAsnProSerLeuHisGluSerHisThrHisThrValGlySerPro 268
QY 824 GATGAGCTGAAGTGTGTTCCACGCGCATTTGCAAGCAACGGTGTCTTCCAGGTGAGC 883
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
269 -----LeuPheAsnGlnValLysIleLysLeuAla----- 279
QY 884 GGGGACTCTGCGCGGAGCAGAAG-----GAGAGCGCAGCAGCCATG 925
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
280 ---AspSerIleLeuGluArgLysArgIleSerAlaAlaArgGluArgLysAlaThrLys 298
QY 926 ATGGTGGGAATTCGATTTGGCTGTTTGTGCTGTGATGCCCTTCTTCTGACGAA 985
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
299 ThrLeuGlyIleIleLeuGlyAlaPheIleIleCysTrpLeuProPhePheValSer 318
QY 986 CTCATCAGCCACTCTGT-----GCCTGACGCTGCCCCCATCTGGAAAGCATATTT 1039
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
319 LeuValLeuProIleCysArgAspSerCysTrpIleHisProAlaLeuPheAspPhePhe 338
QY 1040 CTGTGGCTGTGCTACTCCAAATTTCTTCAACCCCTGATTACAGCTTTTAAACAAG 1099
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
339 ThrTrpLeuGlyTyrLeuAsnSerLeuIleAsnProValIleTyrThrValPheAsnGlu 358
QY 1100 AACTACAAATGCTTCAAGAGCGCTC 1126
Db |||:||||| |||:||||| |||:||||| |||:||||| |||:||||| |||:|||||
```



```
Db 359 AspPheArgGlnAlaPheGlnLysVal 367
RESULT 15
QBGS4
ID QBGS4 PRELIMINARY; PRT; 421 AA.
AC QBGS4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 5-hydroxytryptamine 1A receptor.
GN HTR1A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Cortex, and Hippocampus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK043668; BAC31611.1; -.
DR EMBL; AK049814; BAC33931.1; -.
DR EMBL; AK049884; BAC33970.1; -.
DR MGD; MGI:96273; Htr1a.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
SQ SEQUENCE 421 AA; 46175 MW; 74F519DF7CC0C7AA CRC64;

Alignment Scores:
Pred. No.: 3,32e-20 Length: 421
Score: 423.00 Matches: 116
Percent Similarity: 45.25% Conservative: 65
Best Local Similarity: 29.00% Mismatches: 123
Query Match: 19.64% Indels: 96
DB: 11 Gaps: 11

US-09-976-782-15 (1-1152) x QBGS4 (1-421)
QY 161 TTCACGGTCTGGTGTGAGCGTGTAGTGTG-----CTGATCGCTCCACTTTCCTG 214
Db 33 PheSerTyrGlnValIleThrSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 52
QY 215 TGGAACTGTGTTCCGTCACCATCCCGGGTCCGTCCTCCACCGCGTGGCGCAT 274
Db 53 GlyAsnAlaCysValValAlaAlaIleAlaLeuGluArgSerLeuGlnAsnValAlaAsn 72
QY 275 AACTTGGTGGCTCGACGCGCTCTCGAGCAACTAGTGGCAGCGTGGCGAGTCCACCG 334
Db 73 TyrLeuIleGlySerLeuAlaValThrAspLeuMetValSerValLeuValLeuPro--- 91
QY 335 AGCTTGGGAGTGTGAGTGTGACCGGCGAGCGTGTGCTGCTGGCGGAGCGCTGTGCCAC 394
Db 92 ---MetAlaAlaTyrGlnValLeuAsnLysTyrThrLeuGlyGlnValThrCysAsp 110
QY 395 GTGTGATCTCTTCGACCGGAGCGTCTGTGTGCTGCCCGCGGCTCGGGAACGTG 454
Db 111 LeuPheIleAlaLeuAspVal-----LeuCysCysThrSerSerIleLeuHisLeu 127
QY 455 GGGGCATCGCCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 514
Db 128 CysAlaIleAlaLeuAspArgTyrTrpAlaIleThrAspProIleAspTyrValAsnLys 147
QY 515 ACCGCGAGCGCGCGCTGCTGTGCTCATGATCGCGCTCGCGCGGCGGCGGCGGCGGCTCATC 574
```

```
Db 148 ArgThrProArgAlaAlaAlaLeuIleSerLeuThrTrpLeuIleGlyPheLeu 167
QY 575 GCCCTCGCGCGCTGCTCTTTCGGCGGCGGAGGTGTGCGACGCTCGGCTCCAGCGCTGC 634
Db 168 SerIleProMetLeuLeuGlyTrpArgThrProGluAspArgSerAsnProAsnGluCys 187
QY 635 CAGGTGAGCGGGAACCTCTATGCGCTTCTCCACCGGCGGCGGCGGCGGCGGCGGCGG 694
Db 188 ThrIleSerLysAspHisGlyTyrThrIleTyrSerThrPheGlyAlaPheIlePro 207
QY 695 CTTGGCTGTGTCGCTTCTACCGGAAGATCTACGAGCGCGCCAAAGTTTCCT----- 748
Db 208 LeuLeuLeuMetLeuValLeuTyrGlyArgIlePheArgAlaAlaArgPheArgIleArg 227
QY 749 -----TTCGCG----- 754
Db 228 LysThrValLysLysValGluLysLysGlyAlaGlyThrSerPheGlyThrSerSerAla 247
QY 755 -----CGCGCG----- 760
Db 248 ProProProLysLysSerLeuAsnGlnProGlySerGlyAspCysArgArgSerAla 267
QY 761 ----CGAGAGCTGTG----- 772
Db 268 GluAsnArgAlaValGlyThrProCysAlaAsnGlyAlaValArgGlnGlyGluAspAsp 287
QY 773 -----CTGCGCTGTG 781
Db 288 AlaThrLeuGluValIleGluValHisArgValGlyAsnSerLysGlyHisLeuProLeu 307
QY 782 CGCGCGCACCATCGAGGTGAGGTCCAAAGTAAAG-----GAACACCTGATGAGGCT 832
Db 308 ProSerGluSerGlyAlaThrSerTyrValProAlaCysLeuGluArgLysAsnGluArg 327
QY 833 GAAGTGTGTTTCACGCGACATTTGCAAGCAACCGTGTCTTCCAGGTGAGCGGCGGACTCC 892
Db 328 -----ThrAlaGluAlaLysArgLysMetAlaLeu----- 337
QY 893 TGGCGGAGCAGCAGAGGAGCGAGCAGCATGATGTTGGGAATCTGTATTGGCGTGT 952
Db 338 -----AlaArgGluArgLysThrValLysThrLeuGlyIleIleMetGlyThrPhe 354
QY 953 GTGCTGTGTGATCCCTCTCTCTGACGGAACATCATCAGCCCACTCTGT-----GCC 1006
Db 355 IleLeuCysTyrLeuProPheIleValAlaLeuValLeuProPheCysGluSerSer 374
QY 1007 TCGACGCTGCCCGCATCTGGAAGACATATTTCTGTGCTGTGCTGTGCTGTGCTGTGCT 1066
Db 375 CysHisMetProGluLeuLeuGlyAlaIleIleAsnTyrLeuGlyTyrSerAsnSerLeu 394
QY 1067 TTCACCGCGCTGATTTACACACTTTTACACAGAACTACAACTGCTTTCAGAGCGCTC 1126
Db 395 LeuAsnProValIleTyrAlaTyrPheAsnLysAspPheGlnAsnAlaPheLysIle 414
```

Search completed: August 31, 2004, 20:21:30
Job time : 186 secs


```

Db      378 AlaPheHisLysLeu 382
|||||
|||||
serotonin receptor 1B - human
A:Alternate names: 5-hydroxytryptamine 1D receptor beta; 5-hydroxytryptamine receptor 1B
JN0268
C:Species: Homo sapiens (man)
C:Date: 17-Aug-1992 #sequence_revision 17-Aug-1992 #text_change 05-Nov-1999
R:Accession: JN0268; A42146; A42490; B4136; I57938
R:Mochizuki, D.; Yuyama, Y.; Tsujita, R.; Komaki, H.; Sagai, H.
Biochem. Biophys. Res. Commun. 185, 517-523, 1992
A:Title: Cloning and expression of the human 5-HT1B-type receptor gene.
A:Reference number: JN0268; MUID:92304270; PMID:1610347
A:Accession: JN0268
A:Molecule type: DNA
A:Residues: 1-390 <MOC>
R:Hamblin, M.W.; Metcalf, M.A.; McGuffin, R.W.; Karpells, S.
Biochem. Biophys. Res. Commun. 184, 752-759, 1992
A:Title: Molecular cloning and functional characterization of a human 5-HT1B serotonin r
A:Reference number: JQ1520; MUID:92246962; PMID:1315531
A:Accession: JQ1520
A:Molecule type: DNA
A:Residues: 1-390 <HAM>
A:Cross-references: GB:M89478
R:Jin, H.; Oksenberg, D.; Ashkenazi, A.; Peroutka, S.J.; Duncan, A.M.V.; Rozmahel, R.; Y
J. Biol. Chem. 267, 5735-5738, 1992
A:Title: Characterization of the human 5-hydroxytryptamine-1B receptor.
A:Reference number: A42146; MUID:92210526; PMID:1348246
A:Accession: A42146
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-390 <JIN>
A:Cross-references: GB:M84896
R:Demchynshyn, L.; Sunahara, R.K.; Miller, K.; Teitler, M.; Hoffman, B.J.; Kennedy, J.L.;
Proc. Natl. Acad. Sci. U.S.A. 89, 5522-5526, 1992
A:Title: A human serotonin 1D receptor variant (SHT1Dbeta) encoded by an intronless gene
A:Reference number: A45261; MUID:92302275; PMID:1351684
A:Accession: A45261
A:Molecule type: DNA
A:Residues: 1-390 <DEM>
A:Cross-references: GB:M75128; NID:G184459; PIDN:AAA58675.1; PID:G184460
R:Levy, F.O.; Gudermann, T.; Perez-Reyes, E.; Birnbaumer, M.; Kaumann, A.J.; Birnbaumer,
J. Biol. Chem. 267, 7553-7562, 1992
A:Title: Molecular cloning of a human serotonin receptor (S12) with a pharmacological p
A:Reference number: A42490; MUID:92218412; PMID:1559993
A:Accession: A42490
A:Molecule type: DNA
A:Residues: 1-390 <LEV>
A:Cross-references: GB:M83180; NID:G184463; PIDN:AAA36029.1; PID:G184464
A:Note: sequence extracted from NCBI backbone (NCBIN:94077, NCBIIP:94080)
A:Note: authors proposed that this is the 5-hydroxytryptamine 1D subtype receptor
R:Weinshank, R.L.; Zgombick, J.M.; Macchi, M.J.; Branchek, T.A.; Hartig, P.R.
Proc. Natl. Acad. Sci. U.S.A. 89, 3630-3634, 1992
A:Title: Human serotonin 1D receptor is encoded by a subfamily of two distinct genes: 5-
A:Reference number: A44136; MUID:92228840; PMID:1565658
A:Accession: B44136
A:Molecule type: mRNA
A:Residues: 1-390 <WEI>
A:Cross-references: GB:M81590; NID:G338025; PIDN:AAA60316.1; PID:G338026
R:Velldman, S.A.; Bienkowski, M.J.
Mol. Pharmacol. 42, 439-444, 1992
A:Title: Cloning and pharmacological characterization of a novel 5-hydroxytryptamine 1D
A:Reference number: I57938; MUID:93024334; PMID:1328844
A:Accession: I57938
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-390 <RES>
A:Cross-references: GB:I09732; NID:G184467; PIDN:AAA36030.1; PID:G184468
C:Comment: This protein mediates the diverse actions of 5-hydroxytryptamine in mammals.
C:Genetics:
A:Gene: GDB:HTR1B; S12; 5-HT1B; HTR1D2; 5-HT1DB
A:Cross-references: GDB:132312; OMIM:182131

```

```

A:Map position: 6q14.3-6q16.3
C:Superfamily: vertebrate rhodopsin
C:Keywords: G protein-coupled receptor; glycoprotein; lipoprotein; neurotransmitter rece
F:50-75/Domain: transmembrane #status predicted <TM1>
F:84-112/Domain: transmembrane #status predicted <TM2>
F:124-145/Domain: transmembrane #status predicted <TM3>
F:166-187/Domain: transmembrane #status predicted <TM4>
F:206-229/Domain: transmembrane #status predicted <TM5>
F:315-340/Domain: transmembrane #status predicted <TM6>
F:353-373/Domain: transmembrane #status predicted <TM7>
F:162/Binding site: carboxylate (Asn) (covalent) #status predicted
F:162/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predict
F:252/Binding site: phosphate (Thr) (covalent) (by protein kinase A) #status predict
F:388/Binding site: palmitate (Cys) (covalent) #status predicted

Alignment Scores:
Pred. No.: 5,86e-22 Length: 390
Score: 448.00 Matches: 127
Percent Similarity: 48.35% Conservative: 63
Best Local Similarity: 32.32% Mismatches: 147
Query Match: 20.80% Indels: 56
DB: Gaps: 13

US-09-976-782-15 (1-1152) x JN0268 (1-390)
QY 59 CCCGAGACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATATCGGTTCGACCCCG 118
|||
|||
Db 10 ProProProAlaGlySerGluThrTrpValProGlnAlaAsnLeuSerAlaPro 29
|||
|||
QY 119 AGC-----GGCGCGCTCTCGCGGCGCGA 142
|||
|||
Db 30 SerGlnAsnCysSerAlaLysAspTyrIleTyrGlnAspSerIleSerLeuPro----- 47
|||
|||
QY 143 GGGCGCGCTCTCTCTCTTACGGTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 202
|||
|||
Db 48 -----TrpLysValLeuValMetLeuAlaLeuIleThrLeu 61
|||
|||
QY 203 GGCACCTTCTCTGGAACCTGCTGCTCGGTCACCATCCCGCGGTCGCTGCCCTCCAC 262
|||
|||
Db 62 AlaThrThrLeuSerAsnAlaPheValIleAlaThrValTyrArgThrArgLysLeuHis 81
|||
|||
QY 263 CGCGTGGCCGACAACTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 322
|||
|||
Db 82 ThrProAlaAsnTyrIleAlaSerLeuAlaValThrAspLeuValSerIleLeu 101
|||
|||
QY 323 GCGATGCCACCGAGCTGGCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 382
|||
|||
Db 102 ValMetPro-----IleSerThrMetTyrThrValThrGlyArgThrThrLeuGlyGln 119
|||
|||
QY 383 AGCGTGTGGCAGCTGTGGATCTCTTCGACCGCGGAGCCTGTGTGTGTGTGTGTGTGT 442
|||
|||
Db 120 ValValCysAspPheThrLeuSerSerAsp-----IleThrCysCysThrAlaSer 136
|||
|||
QY 443 CTCGGGAACGTGGCGCCATCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 502
|||
|||
Db 137 IleLeuHisLeuCysValIleAlaLeuAspArgTyrTrpAlaIleThrAspAlaValGlu 156
|||
|||
QY 503 CACAGCTGGCAGCCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 562
|||
|||
Db 157 TyrSerAlaLysArgThrProLysArgAlaAlaValMetIleAlaLeuValTrpValPhe 176
|||
|||
QY 563 TCGGGCGCTATCGCCCTCGCGCGCTGTCTTTTGGCGGGGCGGAGGTGTGCGACGCTCG 622
|||
|||
Db 177 SerIleSerLeuProPro---PhePheThrArgGlnAlaLysAlaGluGluGlu 195
|||
|||
QY 623 CTCACGCGTGCAGGTGAGCGCGGGAACCC---TCTATGCGCGCTCTTCCACCGCGGC 679
|||
|||
Db 196 ValSerGluCysValValAsnThrAspHisIleLeuTyrThrValTyrSerThrValGly 215
|||
|||
QY 680 GCCTTCCACTGCGCTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 739
|||
|||
Db 216 AlaPheTyrPheProThrLeuLeuIleAlaLeuTyrGlyArgIleTyrValGluAla 235
|||
|||

```

```

QY 740 AAGTTTCGTTTC-----GGCCGCCGCG---CGGAGAGCTGTG 772
Db 236 ArgSerArgIleLeuLysGlnThrProAsnArgThrGlyLysArgLeuThrArgAlaGln 255
QY 773 CTG-----CCGTGCGCGCCACCATGTCAGGTGAGGTCCAAAGTAAG 814
Db 256 LeuIleThrAspSerProGlySerThrSerSerValThrSerIleAsnSerArgValPro 275
QY 815 GAAGCACCCTGATGAGCTT---GAAGTGGTGTTCACCGGCACATTGCAAAAGCAACGTGTCC 871
Db 276 AspValProSerGluSerGlySerProValTyrrValAsnGlnValLysValArgValSer 295
QY 872 TTCAGGTGAGCGGGACTCTCTGCGCGGAGCAGAAG-----GAGAGG 913
Db 296 -----AspAlaLeuLeuGluLysLysLysLeuMetAlaAlaArgGluArg 310
QY 914 CGAGCAGCCATGATGGTGGGAATCTGATTTGGCGTGTGTGCTGTGTCGTGATCCCTTC 973
Db 311 LysAlaThrLysThrLeuGlyIleLeuGlyAlaPheIleValCysTrpLeuProPhe 330
QY 974 TTCCTGACGGAACTCATCAGCCCATCTGT-----GCTGCAGCCTGCCCCCATCTGG 1027
Db 331 PheIleIleSerLeuValMetProIleCysLysAspAlaCysTrpPheHisLeuAlaIle 350
QY 1028 AAAAGCATATTTCTGTGCTGGTCTGCTACCTCAATCTTTCTCAACCCCTGATTTACACA 1087
Db 351 PheAspPhePheThrTrpLeuGlyTyrrLeuAsnSerLeuIleAsnProIleIleTyrr 370
QY 1088 GCTTTTACAGAACTACAAACATGCTTCAAGACCTC 1126
Db 371 MetSerAsnGluAspPheGlnAlaPheHisLysLeu 383

RESULT 8
S58126
serotonin receptor 1-like - rabbit
N;Alternate names: 5-hydroxytryptamine receptor 1-like (5-HTR)
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 13-Jan-1996 #sequence_revision 01-Mar-1996 #text_change 05-Nov-1999
C;Accession: S58126
R;Wurch, T.; Cathala, C.; Palmer, C.; Valentin, J.P.; John, G.; Colpaert, F.C.; Pauwels,
submitted to the EMBL Data Library, July 1995
A;Reference number: S58126
A;Accession: S58126
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-390 <WUR>
A;Cross-references: EMBL:X99731; NID:g1052771; PIDN:CAA61893.1; PID:g1052772
C;Superfamily: vertebrate rhodopsin
C;Keywords: neurotransmitter receptor

Alignment Scores:
Pred. No.: 1.08e-21 Length: 390
Score: 444.00 Matches: 117
Percent Similarity: 51.30% Conservative: 60
Best Local Similarity: 33.91% Mismatches: 132
Query Match: 20.61% Indels: 36
DB: 2 Gaps: 11

US-09-976-782-15 (1-1152) x S58126 (1-390)

QY 167 GTCTGTGTGAGCGTGTAGTCTGCTGATCGCTGCCACTTTCCTGTGGAACCTGCTG 226
Db 50 ValLeuLeuValLeuLeuAlaLeuPheThrLeuAlaThrThrLeuSerAsnAlaPhe 69
QY 227 GTTCCGGTCCACCATCCCGGGTCCGCTTCCACCGGTGCGCATACTTGTGGCC 286
Db 70 ValValAlaThrValTyrrArgThrArgLysLeuHisThrProAlaAsnTyrrLeuAla 89
QY 287 TCGACGGCCCTCTCGGACGAACTAGTGGCAGCTGTGGGATGCCACCGAGCTGGCGAGT 346
Db 90 SerLeuAlaValThrAspLeuValSerIleLeuValMetPro-----IleSerThr 107
QY 347 GAGCTGTGACCGGCGACGTGGCTGCTGGCGCGGAGCCTGTGCCACGTGTGATCTCC 406

```

```

Db 108 MetTyrrThrValThrGlyArgTrpThrLeuGlyGlnValValCysAspLeuTrpLeuSer 127
QY 407 TTCGACCGCGGAGCTGTCTGTGTCGCCCGCGCTCGGGAACGTGGCGCCATCGCC 466
Db 128 SerAsp-----IleThrCysCysThrAlaSerIleMetHisLysValIleAla 144
QY 467 CTGGGCGCGGACCGGCGCATCACACGGCACCTGCGAGCACCGCTGCGGACCGCCG 526
Db 145 LeuAspArgTyrrTrpAlaIleThrAspAlaValGluTyrrSerAlaLysArgThrProLys 164
QY 527 GCCTGTGCTCATCATCGCTCGCCCGGGTGGCGCTCGGCGCTCATGCCCTCGCGCG 586
Db 165 ArgAlaAlaIleMetIleArgLeuValTrpValPheSerIleCysIleSerLeuPro 184
QY 587 CTGCTCTTTGGCGGCGGAGGTGTGCGACGCTCGGCTCCAGCGCTGCGAGGTGAGCGG 646
Db 185 ---PhePheTrpArgGlnAlaLysAlaGluGluValSerGluCysLeuValAsnThr 203
QY 647 GAA---CCCTCTATGCGCGCTTCTCCACCGCGCGGCTTCCACCTGCGCTTGGCGTG 703
Db 204 AspHisValLeuTyrrThrValTyrrSerThrValGlyAlaPheTyrrLeuProThrLeu 223
QY 704 GTGCGCTTTGTCTACCGGAGATCTACGAGCGCGCAAGTTTCGTTTC----- 751
Db 224 LeuIleAlaLeuTyrrGlyArgIleTyrrValGluAlaArgSerArgIleLeuLysGlnThr 243
QY 752 -----GGCGCGCGC---CGGAGAGCTGTGCTG-----CCG 778
Db 244 ProAsnArgThrGlyLysArgLeuThrArgAlaGlnLeuIleThrAspSerProGlySer 263
QY 779 TTGCGCGCCACCATGACGAGTCCAGGTGCTCCAAAGTAAAGNAGCACCTGATCAGCT 835
Db 264 ThrThrSerValThrSerIleAsnSerArgAlaProAspValProSerGluSerGlySer 283
QY 836 GTGCTGTTCACGGCACATTGCAAAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCTGG 895
Db 284 ProValTyrrValAsnGlnValLysValArgValSer-----AspAlaLeu 298
QY 896 CGGAGCAGCAGAG-----GAGAGCGGAGCAGCCATGATGTGGGAATT 937
Db 299 LeuGluLysLysLysLeuMetAlaAlaArgGluArgLysAlaThrLysThrLeuGlyIle 318
QY 938 CTGATTGCGCGTGTGTGCTGTGCTGATCCCTTCTTCTTCGCGGAACTCATCAGCCCA 997
Db 319 IleLeuGlyValPheIleValCysTrpLeuProPhePheIleIleSerLeuValMetPro 338
QY 998 CTCTGT-----GCTGTGAGCCTGCCCGCCCATCTCGAAAGACATATTCTGTGCTGGC 1051
Db 339 IleCysLysAspAlaCysTrpPheHisGlnAlaIlePheAspPhePheThrTrpLeuGly 358
QY 1052 TACTCCCAATTTCTTCTCAACCCCTGATTTACAGCTTTTAAACAAGAACTACAACAAT 1111
Db 359 TyrValAsnSerLeuIleAsnProIleIleTyrrThrMetSerAsnGluAspPheLysGln 378
QY 1112 GCCTTCAAGAGCCTC 1126
Db 379 AlaPheHisLysLeu 383

RESULT 9
B30341
G protein-coupled receptor RDC4 - dog
C;Species: Canis lupus familiaris (dog)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 05-Nov-1999
C;Accession: B30341; S12822
R;Libert, F.; Parmentier, M.; Lefort, A.; Dinsart, C.; Van Sande, J.; Maenhaut, C.; Simor
Science 244, 569-572, 1989
A;Title: Selective amplification and cloning of four new members of the G protein-coupled
A;Reference number: A30341; MUID:89242119; PMID:2541503
A;Accession: B30341
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-377 <LIB>

```


C;Species: Drosophila melanogaster
 C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 04-Sep-1998
 C;Accession: A38271
 R;Witz, P.; Amlaiky, N.; Plassat, J.L.; Maroteaux, L.; Borrelli, E.; Hen, R.
 Proc. Natl. Acad. Sci. U.S.A. 87, 8940-8944, 1990
 A;Title: Cloning and characterization of a Drosophila serotonin receptor that activates
 A;Reference number: A38271; MUID:91062395; PMID:2174167
 A;Accession: A38271
 A;Molecule type: mRNA
 A;Residues: 1-564 <WIT>
 A;Cross-references: GB:M55533
 A;Note: the authors translated the codon CTT for residue 213 as Ala, GTG for residue 215
 C;Genetics:
 A;Gene: FlyBase:5-HT7
 A;Cross-references: FlyBase:FBgn0004573
 C;Superfamily: octopamine receptor type I
 C;Keywords: G protein-coupled receptor; glycoprotein; membrane protein; phosphoprotein

Alignment Scores:
 Pred. No.: 3,88e-21 Length: 564
 Score: 435.50 Matches: 136
 Percent Similarity: 45.15% Conservative: 69
 Best Local Similarity: 29.96% Mismatches: 153
 Query Match: 20.22% Indels: 96
 DB: 2 Gaps: 12

US-09-976-782-15 (1-1152) x A38271 (1-564)

QY	14	GCTAGCCTTTCAGTGGCCACCGCGGCTTCCCTTGGCCCTGGGACCCGAGACACGACG 73
DB	83	AlaSerAlaSerSerThrGlySerGlySerGlySerGlySerGlySerGlySer 102
QY	74	GGACCCGGGACC-----CCAGCCCGAGAGGATCTCGT--- 109
DB	103	GlySerGlySerThrGlyLeuAlaSerMetAsnSerSerProIleAlaIleValSerTyr 122
QY	110	-----TGACCCCGAGCGCC-----GCGTCTCTCCGCGCGCGAGG 145
DB	123	GlnGlyIleThrSerThrAsnLeuGlyAspSerAsnThrThrLeuValProLeuSerAsp 142
QY	146	CCGCGCTTC-----TCGTCTTCACG 166
DB	143	ThrProLeuLeuLeuGluGluPheAlaAlaGlyGluPheValLeuArgProLeuThrSer 162
QY	167	GTCCTGGTGGTACGCTCTAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 226
DB	163	IlePheValSerIleValLeuLeuIleValIleLeuGlyThrValValGlyAsnValLeu 182
QY	227	GTTCCGGTACATCCCGCGGTCGTCGTCCTTCCACCGCGTCCGACATCTGTGTGGCC 286
DB	183	ValCysIleAlaValCysMetValArgLysLeuArgProCysAsnTyrLeuVal 202
QY	287	TCGACGGCGTCTGGAGCACTAGTGGCAGCGTGGCGATGCCACGAGCTGGCGAGT 346
DB	203	SerLeuAlaLeuSerAspLeuCysValAlaLeuValMetProMetAlaLeuTyr 222
QY	347	GAGCTGTCCGCGGCGAGTGGTGTGTGGCGGAGCTGTGCCAGCTGTGGATCTCC 406
DB	223	GluValLeu-----GluYtrtrPAsnPheGlyProLeuLeuCysAspIleTyrValSer 240
QY	407	TTCGACCGCGGAGCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 466
DB	241	PheAspVal-----LeuCysCysThrAlaSerIleLeuAsnLeuCysAlaIleSer 257
QY	467	CTGGGCGCGGCGGCGCATCACAGGCCTGTGACGACACAGCTGTGCGCAGCCGCGCGCG 526
DB	258	ValAspArgTyrLeuAlaIleThrIlyProLeuGluTyrGlyValIlyAspThrProArg 277
QY	527	GCCTGTGTGTCATGATCGCTCGCCCGGTCGCGTGGCGCTCATCGCTCGCGCG 586
DB	278	ArgMetMetLeuCysValGlyIleValTrpLeuAlaAlaIleCysIleSerLeuPro 297
QY	587	CTGCTCTTTGGCGGGGCGAGGTGTGCGACGCTCGGCTCCAGCGC---TGCGAGGTGAGC 643

DB	238	LeuLeuIleLeuGlyAsnGluHisGluAspGluGluGlyGlnProIleCysThrValCys 317
QY	644	CGGGAACCTCTATGCGCGCTTCTCCACCGCGCGCGCTTCCACCTGCGCTTGGCGTG 703
DB	318	GlnAsnPheAlaTyrGlnIleTyrAlaThrLeuGlySerPheTyrIleProLeuSerVal 337
QY	704	GTGCGGTTTGTCTACCGGAAGATCTACGAGGCGGCAAGTTTCGTTTCGCGCGCGCGG 763
DB	338	MetLeuPheValTyrTyrGlnIlePheArgAlaAla-----Arg 350
QY	764	AGAGCTGTGCTG----- 775
DB	351	ArgIleValLeuGluGluLysArgAlaGlnThrHisLeuGlnGlnAlaLeuAsnGlyThr 370
QY	776	-----CCGTGCGCGCCACCATGCAA-----AGCAACGGTGTCTCTCCAGGTG 808
DB	371	GlySerProSerAlaProGlnAlaProLeu--GlyHisThrGluLeuAlaSerSerG 390
QY	809	GTAAAGGA-----AGCACTGATGAG 829
DB	390	LysGlyGlnArgHisSerValGlyAsnThrSerLeuThrTyrSerThrCysGlyG 410
QY	830	GCTGAAGTGGTGTTCACCGCACATGCAA-----AGCAACGGTGTCTCTCCAGGTG 880
DB	410	LysLeuSerSerGlyGlyAlaLeuAlaGlyHisGlySerGlyGlyValSerGlyS 430
QY	881	AGCGGGGACTCTCTGCGGGA-----CGAGAGGAG 910
DB	430	erThrGlyLeuLeuGlySerProHisLysLysLeuArgPheGlnLeuAla--LysGlu 449
QY	911	AGCGGAGCAGCATCATGTCGGGAATCTGATGCGCGTGTGTGTGTGTGTGTGTGTGT 970
DB	450	LysLysAlaSerThrThrLeuGlyIleIleMetSerAlaPheThrValCysTrpLeuPro 469
QY	971	TCTCTCTGACGGAACATCATCAGCCCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1030
DB	470	PhePheIleLeuAlaLeuIleArgProPheGluThrMetHisValProAlaSerLeuSer 489
QY	1031	AGCATATTCTGT 1090
DB	490	SerLeuPheLeuTyrLeuGlyTyrAlaAsnSerLeuLeuAsnProIleIleTyrAlaThr 509
QY	1091	TTTACAGAACTACAAATGCTTCAAGAGCTC 1126
DB	510	LeuAsnArgAspPheArgLysProPheGlnGluIle 521

RESULT 14
 T42203
 serotonin receptor 5-HT precursor - Caenorhabditis elegans
 C;Species: Caenorhabditis elegans
 C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jan-2000
 C;Accession: T42203
 R;Olde, B.; McCombie, W.R.
 J. Mol. Neurosci. 8, 53-62, 1997
 A;Title: Molecular cloning and functional expression of a serotonin receptor from Caenorh
 A;Reference number: 222071; MUID:97215250; PMID:9061615
 A;Accession: T42203
 A;Status: preliminary; translated from GB/EMBL/DBJ
 A;Molecule type: mRNA
 A;Residues: 1-445 <OLD>
 A;Cross-references: EMBL:U15167; PIDN:AA866360.1
 A;Experimental source: clone CER-1
 C;Superfamily: vertebrate rhodopsin
 C;Keywords: transmembrane protein

Alignment Scores:
 Pred. No.: 7.14e-21 Length: 445
 Score: 431.50 Matches: 135
 Percent Similarity: 44.24% Conservative: 61
 Best Local Similarity: 30.47% Mismatches: 156
 Query Match: 20.03% Indels: 91
 DB: 2 Gaps: 13

US-09-976-782-15 (1-1152) x T42203 (1-445)

Qy	47	CTTGCCCTGGACCGGACCAGCAGCAGCGGACCCGGGACCCCAAGCCCGGAGAGGATACTC	106
Db	7	LeuAsnLeuThrAlaSerProSerGlnProAlaThr-----LeuLeu 21	
Qy	107	GGTTGCACCCGAGCGCGCGTCCGTG-----CCGGGCGAGGGCGGCCCTTC 154	
Db	22	SerAlaValAlaArgGlyThrHisLeuValAspGlnPheProAlaHisAlaGluIlePhe 41	
Qy	155	TCT-----GTCTTCACGTCCTGGTGGTGAACGCTGCTACTGTCTGTCTGCTG 196	
Db	42	SerAspIleGluArgProLeuValGlnThrValIleLeuAlaSerValLeuValLeu 61	
Qy	197	ATCGCTGCACATTCTCTGTGGAACTGCTGTTCCTGGTCACCATCCGCGGTCCGTGCG 256	
Db	62	IleLeuSerCysPheIleGlyAsnLeuPheValIleLeuAlaIleMetGluArgAsp 81	
Qy	257	TTCCAC-----CGGTCGCCGATNACTGTGTGGCTCGACGCGCGTCTCGGACGAAC TAGTG 313	
Db	82	LeuArgGlyArgProGlnIntyrTrpLeuIlePheSerLeuAlaValAlaAspLeuVal 101	
Qy	314	GCAGCGCTGGGATGCCACGAGCTGGCGAGTAGCTGTCACCGGGCGAGCTCGGCTG 373	
Db	102	GlyMetIleValThrPro-----LeuGlyAlaTrpPheThrValThrGlyThrTrpAsn 119	
Qy	374	CTGGGCGGAGCTGTGCCACGTGTGGATCTCTTCGACGCGGAGCCTGTCTGTGCTGC 433	
Db	120	LeuGlyValValValCysAspPheTrpIleSerValAspVal-----LeuValCys 136	
Qy	434	CCGCGCGGCTCGGAAAGTGGCGGCACATCGCCCTGGCGCGACGGGCGCCATCACACG 493	
Db	137	ThrAlaSerIleLeuHisLeuValAlaIleAlaLeuAspArgTrpSerIleThrAsp 156	
Qy	494	CACCTGCACACACGCTGGCGCACCGCGACCGCGCTCGTGTCTCATGATCGGCTCGCG 553	
Db	157	IleCysTyrrValGlnAsnArgThrProLysArgIleThr--LeuMetLeuAlaIle 175	
Qy	554	CGGTGCGCTCGGCGCTCATCGCCCTCGCGCGCTGCTCTTTGGCGGGCGAGGTGTCG 613	
Db	176	TrpPheThrSerLeuLeuIleSerLeuAlaProPheAlaGlyTrpLysAspGluGlyPhe 195	
Qy	614	GACGCTCGGCTCCAGCGC-----TCCAGGTGAGCGCGGAACCTCCATTCGCGCG 664	
Db	196	SerAspArgValLeuLysSerHisValCysLeuIleSerGlnIleSeryrGlnVal 215	
Qy	665	TTCTCCACCGCGCGCCTTCACCTCGCGCTTGGCGTGGTCCGCTTGTCTACCGAAG 724	
Db	216	PheSerThrAlaThrAlaPheTyrrIleProLeuIleAlaIleCysValTyrrTrpLys 235	
Qy	725	ATCTACGAGCGCGCAAAGTTTCGTTTCGCGCGC-----CGCCGGAGAGCTGTGCTG 775	
Db	236	IleMetArgAlaAlaLysLysArgPheLysArgGluArgAspArgThrValIleArg 255	
Qy	776	CGTTTGCGC-----GCCACCATCGAGTGATGCCAAG-----808	
Db	256	ProProProAspAlaIleAspGluLysLysAlaMetMetProLysLysSerLysLysCys 275	
Qy	808	-----808	
Db	276	ProLeuProProAlaValValIleSerAspIleGlnAlaAsnGlyGlyThrGlyGlyLys 295	
Qy	809	-----GTAAGAAGAACACCT-----GAT 826	
Db	296	ThrAsnSerIleLysAsnProProArgHisAsnGluSerSerSerAlaSerGluGlu 315	
Qy	827	GAGGCTGAAGTGGTGTTCAGGCACATTGCCAAGCAACGCTGCTTCCTCAGGTG-----880	
Db	316	GluArgThrMetThrGlnThrAsnHisAlaProGlyAspValThrGlnGluThrLys 335	
Qy	881	-----ACGCGG 886	

RESULT 15

A53279

serotonin receptor 1D - human

N/Alternate names: 5-hydroxytryptamine receptor 1D (5-HT1D); 5HT-1D alpha receptor

C/Species: Homo sapiens (man)

C/Date: 12-May-1994 #sequence_revision 12-May-1994 #text_change 05-Nov-1999

C/Accession: A53279; A44136

R/Hamblin, M.W.; Metcalf, M.A.

Mol. Pharmacol. 40, 143-148, 1991

A/Title: Primary structure and functional characterization of a human 5-HT-1D-type sero

A/Reference number: A53279; MUID:91342595; PMID:1652050

A/Accession: A53279

A/Molecule type: DNA

A/Residues: 1-377 <HAM>

A/Cross-references: GB:M89955; NID:G177771; PIDN:AAA35491.1; PID:G177772

R/Weinshank, R.L.; Zgombick, J.M.; Macchi, M.J.; Branchek, T.A.; Hartig, P.R.

Proc. Natl. Acad. Sci. U.S.A. 89, 3630-3634, 1992

A/Title: Human serotonin 1D receptor is encoded by a subfamily of two distinct genes: 5-

A/Reference number: A44136; MUID:92228840; PMID:1565658

A/Accession: A44136

A/Molecule type: mRNA

A/Residues: 1-377 <WEI>

A/Cross-references: GB:M81589; NID:G338023; PIDN:AAA60315.1; PID:G338024

C/Genetics:

A/Gene: GDB:HTR1D

A/Cross-references: GDB:132416; OMIM:182133

A/Map position: lp36.3-lp34.3

A/Introns: #status absent

C/Superfamily: octopamine receptor type I

C/Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transme

Alignment Scores:	
Pred. No.:	Length:
Score:	377
Percent Similarity:	Matches: 118
Best Local Similarity:	Conservative: 70
Query Match:	Mismatches: 156
DB:	Indels: 40
	Gaps: 12

US-09-976-782-15 (1-1152) x A53279 (1-377)

Qy	53	CTGGGACCGAGCAGCAGCGGACCGGACCGCCA-----AGCCGAGAGGGATA	103
	:::		:::
Db	1	MetSerProLeuAsnGlnSerAlaGluGlyLeuProGlnGluAlaSerAsnArgSerLeu	20
	:::		:::
Qy	104	CTCGGTTGCAACCGGCGGCGGCGTCTCTCCGCGCGGAGCGCGCCCTCTCTGCTCTC	163
	:::		:::
Db	21	AsnAlaThrGluThrSerGluAlaTrpAspProArgThrLeuGlnAlaLeuLysIleSer	40
	:::		:::
Qy	164	ACGGTCTGTGTGAGCGCTGTAGTCTGTATGCTGCTGACCTGTCACCTTTCCTGTGGAACCTG	223
	:::		:::

```
Db      41  LeuAlaValValLeuSerValIleThrLeu-----AlaThrValLeuSerAsnAla 57
QY      224 CTGTTCCGGTCACATCCCGGGTCGGTCCCTTCACCGGTCGCCATAACTTGGTG 283
Db      58  PheValLeuThrIleLeuLeuThrArgLysLeuHisThrProAlaAsnTyrLeuIle 77
QY      284 GCCTCGACGGCGCTCTCGAGAACTAGTGGCAGCGCTGGGGATGCCACCGAGCCTGGCG 343
Db      78  GlySerLeuAlaThrThrAspLeuValSerIleLeuValMetProIleSerIleAla 97
QY      344 AGTGAGCTGTGCACGGCGCAGCTGGCTGTGGCCGGAGCTGTGCCACGTGTGGATC 403
Db      98  TyrThrIleThr-----HisThrTrpAsnPheGlyGlnIleLeuCysAspIleTrpLeu 115
QY      404 TCCITCGACGGCGGAGCTGTGTGTCGCCCGCGGCTCGGAAAGTGGCGGCCATC 463
Db      116 SerSerAsp-----IleThrCysThrAlaSerIleLeuHisLeuCysValIle 132
QY      464 GCCCTGGCGCGCGGCCCATCACACGGCACCTGCACACAGCTGCGCACCGCCGACG 523
Db      133 AlaLeuAspArgTyrTrpAlaIleThrAspAlaLeuGluTyrSerLysArgThrAla 152
QY      524 CGCGCTCGTGTCTCATCATCGCTCGCCCGGGTCCGTCGGCGCTCATCGCCCTCGCG 583
Db      153 GlyHisAlaAlaThrMetIleAlaIleValTrpAlaIleSerIleCysIleSerIlePro 172
QY      584 CCGCTGTCTTTTGGCGGGCGGAGGTGTCGACGCTCGGCTCCAGCGCTGCAGGTG--- 640
Db      173 Pro---LeuPheTrpArgGlnAlaLysAlaGlnGluMetSerAspCysLeuValAsn 191
QY      641 AGCGGGAAACCTCTATGCGCGCTCTTCCACCGCGCGCTTCCACCTGCGCGTTCGGC 700
Db      192 ThrSerGlnIleSerTyrThrIleTyrSerThrCysGlyAlaPheTyrIleProSerVal 211
QY      701 GTGTGCGGTTTGTCTACCGAAGATCTACGAGCGGCGCAAGTTTCGT----- 748
Db      212 LeuLeuIleIleLeuTyrGlyArgIleTyrArgAlaAlaArgAsnArgIleLeuAsnPro 231
QY      749 -----TTCGGCGCGCGGAGAGCTGTGTCGCCGTTCGCGGCCACC----- 790
Db      232 ProSerLeuTyrGlyLysArgPheThrThrAlaHisLeuIleThrGlySerAlaGlySer 251
QY      791 -----ATGACAGTGAGTCCCAAGTAAAGAACACCTGATGAGCT-----GAAGTG 838
Db      252 SerLeuCysSerLeuAsnSerSerLeuHisGluGlyHisSerHisSerAlaGlySerPro 271
QY      839 GTGTTTCAGGCACATTGCAAGCAACGGTGTCTCTTCAGGTGAGCGGGGACTCTCTGGCGG 898
Db      272 LeuPhePheAsnHisValLysIleLysLeuAla-----AspSerAlaLeu 286
QY      899 GAGCAGAG-----GAGAGCGAGCAGCCATCATGCTGGGAATTCGT 940
Db      287 GluArgLysArgIleSerAlaAlaArgGluArgLysAlaThrLysIleLeuGlyIleIle 306
QY      941 ATTGCGGTGTTGTGCTGTGTGATCCCTTCTTCAGCGGAACCTCATCAGCCCACTC 1000
Db      307 LeuGlyAlaPheIleIleCysTrpLeuProPhePheValSerLeuValLeuProIle 326
QY      1001 TGT-----GCCTCGAGCTCCCGCCCATCTGGAAAGCATATTTCTGTGCTGGCTAC 1054
Db      327 CysArgAspSerCysTrpIleHisProAlaLeuPheAspPhePheThrTrpLeuGlyTyr 346
QY      1055 TCCATTCTTTTCAACCCCTGATTTACACAGCTTTTACAGAACTACACAAATGCC 1114
Db      347 LeuAsnSerLeuIleAsnProIleIleTyrThrValPheAsnGluPheArgGlnAla 366
QY      1115 TTCAAGAGCCTC 1126
Db      367 PheGlnLysIle 370
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: August 31, 2004, 20:21:44 ; Search time 160 Seconds

(without alignments)
4530.426 Million cell updates/sec

Title: US-09-976-782-15

Perfect score: 2154

Sequence: 1 cgcacgagccgctagcc.....aagcagagatgaacacaggg 1152

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1297172 seqs, 314612898 residues

Total number of hits satisfying chosen parameters: 2594344

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q=/cgn2_1/USPRO_spool_p/US09976782/runat_31082004_115953_10443/app_query.fasta_1.1351
-DB=Published Applications AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09976782@cgn_1_1_225_@runat_31082004_115953_10443
-NCPU=6 -ICPU=3 -NO MAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications AA:

1: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubppa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pap.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubppa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubppa/US09A_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubppa/US09B_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubppa/US09C_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pap.*
13: /cgn2_6/ptodata/2/pubppa/US10A_PUBCOMB.pap.*
14: /cgn2_6/ptodata/2/pubppa/US10B_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubppa/US10C_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pap.*
17: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pap.*
18: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	ID	Description

ALIGNMENTS

RESULT 1

US-09-976-782-16
; Sequence 16, Application US/09976782
; Publication No. US20030190715A1
; GENERAL INFORMATION:
; APPLICANT: Grosse et al
; TITLE OF INVENTION: No. US20030190715A1e1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-157
; CURRENT APPLICATION NUMBER: US/09/976,782
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,113
; PRIOR FILING DATE: 2000-10-12
; PRIOR APPLICATION NUMBER: 60/240,662
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,732
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,625
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,703
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/241,190
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/240,637
; PRIOR FILING DATE: 2000-10-16

Sequence 16, Appl
Sequence 15, Appl
Sequence 10, Appl
Sequence 12, Appl
Sequence 8, Appl
Sequence 30, Appl
Sequence 56, Appl
Sequence 57, Appl
Sequence 56, Appl
Sequence 101, Appl
Sequence 58, Appl
Sequence 57, Appl
Sequence 102, Appl
Sequence 2, Appl
Sequence 59, Appl
Sequence 103, Appl
Sequence 698, Appl
Sequence 610, Appl
Sequence 60, Appl
Sequence 58, Appl
Sequence 104, Appl
Sequence 446, Appl
Sequence 38, Appl
Sequence 16, Appl
Sequence 447, Appl
Sequence 60, Appl
Sequence 59, Appl
Sequence 105, Appl
Sequence 18, Appl
Sequence 17, Appl
Sequence 4, Appl
Sequence 4, Appl
Sequence 15, Appl
Sequence 20, Appl
Sequence 40, Appl
Sequence 427, Appl
Sequence 6, Appl
Sequence 2, Appl
Sequence 6, Appl
Sequence 14, Appl
Sequence 19, Appl
Sequence 3, Appl
Sequence 13, Appl
Sequence 429, Appl
Sequence 10, Appl

```
; PRIOR APPLICATION NUMBER: 60/240,669
; PRIOR FILING DATE: 2000-10-16
; PRIOR APPLICATION NUMBER: 60/262,455
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: 60/240,648
; PRIOR FILING DATE: 2000-10-16
; NUMBER OF SEQ ID NOS: 127
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 16
; LENGTH: 379
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-976-782-15

Alignment Scores:
Pred. No.: 1,126-112 Length: 379
Score: 1946.00 Matches: 379
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 90.34% Indels: 0
Dbs: 10 Gaps: 0

US-09-976-782-15 (1-1152) x US-09-976-782-16 (1-379)

QY 5 ATGGAGGCGCTAGCTTTTCAGTGGCCACCGCGCGGTTGCCCTTGGCCCTGGGACCCGAG 64
Db 1 MetGluAlaAlaSerLeuSerValAlaThrAlaGlyValAlaLeuAlaLeuGlyProGlu 20

QY 65 ACAGAGGCGGACCGCGGACCCCAAGCCCGAGAGGAGTACTCGGTTCGACCCCGAGCGGC 124
Db 21 ThrSerSerGlyProGlyThrProSerProArgGlyIleLeuGlySerThrProSerGly 40

QY 125 GCGGCTCTGCGGCGGCGGCGGCGCTCTCTGCTTCCACGCTCTCTGCTGCTGCTGCTG 184
Db 41 AlaValLeuProGlyArgGlyProProPheSerValPheThrValLeuValValThrLeu 60

QY 185 CTAGTGTCTGCTGATCGCTGCCACTTTCTGTGGAACTCTGCTGCTGCTGCTGCTGCTG 244
Db 61 LeuValLeuLeuIleAlaAlaThrPheLeuTyrAsnLeuLeuValProValThrIlePro 80

QY 245 CGGTGCTGCTTTCACCGGCTGCGGATACCTTGGTGGCTCGACGCGCGCTCTCGGAC 304
Db 81 ArgValArgAlaPheHisArgValProHisAsnLeuAlaSerThrAlaValSerAsp 100

QY 305 GAACCTAGTGGGACGCTGCGGATGCCACGAGCTGCGGAGTGAGCTGCGACCGGCGCA 364
Db 101 GluLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGluLeuSerThrGlyArg 120

QY 365 CGTGGCTGCTGCGGCGGAGCTGTGCCACGCTGTGGATCTCTCTGACCGCGGAGCTGT 424
Db 121 ArgArgLeuLeuGlyArgSerLeuCysHisValTyrPheSerPheAspAlaGlyAlaCys 140

QY 425 CTGTGCTGCGGCGGCTCGGAACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 484
Db 141 LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla 160

QY 485 ATCACAGGACCTGTCAGCACAGCTGCGCACCGCGGCGGCGGCGGCTGTGTCTCATGATC 544
Db 161 IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle 180

QY 545 GCGCTCGCGGCGGCTGCTGCGGCGGCTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCG 604
Db 181 AlaLeuAlaArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly 200

QY 605 GAGGTGTGGAGCTGCGGCTCGAGGCTGCGGAGTGGCGGGAACCTCTCTATGCGGCG 664
Db 201 GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerThrAlaAla 220

QY 665 TTCTCCACCGCGGCGGCTTCCACCTGCGGCTGCGGCTGCGGCTGCGGCTGCTACCGGAAG 724
Db 221 PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValThrArgLys 240

QY 725 ATCTACGAGGCGGCCAAGTTTTCGTCGCGCGCGCGGAGAGCTGTGTCGCGTTCGCGG 784
```

```
Db 241 IleTyrGluAlaAlaLysPheArgPheGlyArgArgAlaValLeuProLeuPro 260
QY 785 GCCACATGCGAGTCCAGGTAAAGGAGGACACCTGATGAGCTCAAGTGGTTC 844
Db 261 AlaThrMetGlnValArgSerLysValLysGluAlaProAspGluAlaGluValPhe 280
QY 845 ACGGCACATTTGCAAGCAACGCTGCTCTTCCAGGTGAGCGGGAGCTCTCTGGCGGAGCAG 904
Db 281 ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln 300
QY 905 AAGGAGGCGAGCGGACCATGATGTTGGAAATCTGATGGCGTGTGTTGCTGCTGCTGG 964
Db 301 LysGluArgArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 320
QY 965 ATCCCTCTCTCTGAGCGGACCTCATCGCCACCTCTGCTGCTGCTGCTGCTGCTGCTGCT 1024
Db 321 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 340
QY 1025 TCGAAAAGCATATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1084
Db 341 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 360
QY 1085 ACAGCTTTTAAACAAGTACAACTGCTTCAAGAGCTCTTCTTACTTAAGCAGAGA 1141
Db 361 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 379
```

RESULT 2

```
US-10-333-946-15
; Sequence 15, Application US/10333946
; Publication NO. US20040023252A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.
; APPLICANT: BURFORD, Neil; YUE, Henry
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.
; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Narinder K.
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.
; APPLICANT: POLJUK, Jennifer L.; KEARNEY, Liam
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.
; APPLICANT: LEE, Ernestine A.; DING, Li
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: PI-0176 USN
; CURRENT APPLICATION NUMBER: US/10/333,946
; PRIOR FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: PCT/US01/23433
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: US 60/221,478
; PRIOR FILING DATE: 2000-07-27
; PRIOR APPLICATION NUMBER: US 60/223,268
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/227,054
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 60/231,121
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: US 60/232,243
; PRIOR FILING DATE: 2000-09-13
; PRIOR APPLICATION NUMBER: US 60/232,691
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/235,146
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7476053CD1
US-10-333-946-15
```


59	LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValProValThrIlePro	78
245	CGGGTCGGTCCTCCACCGCGCGCAATACTTGGTCGCCTCGACGCGCGCTCTCGGAC	304
79	ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp	98
305	GAACCTAGTCGACGGCTGGCGATGCCACCGAGCCTCGCGAGTGAGTGTGACCGGGCGA	364
99	GlulLeuValAlaAlaLeuAlaMetProProSerLeuAlaSerGlulLeuSerThrGlyArg	118
365	CGTCGGCTGTGGCGCGGACCTGTGCCACGTGTGGATCTCTCTTCACGCCGAGCGCTGT	424
119	ArgArgLeuLeuGlyArgSerLeuCysHisValTrpIleSerPheHisVal	135
425	CTGTGTCGCCCGCGCGCTCGGAAACGTGGCGGCATCGCCCTGGCGCGCGACGCGGCC	484
136	LeuCysCysProAlaGlyLeuGlyAsnValAlaAlaIleAlaLeuGlyArgAspGlyAla	155
485	ATCACCGGCACCTGCAGCACAGCTCGGCACCGCAGCGCGCGCTGTGTCTCATGATC	544
156	IleThrArgHisLeuGlnHisThrLeuArgThrArgSerArgAlaSerLeuLeuMetIle	175
545	GCGCTCGCCGGTGGCGTGGCGCTCATCGCCCTCGCGCGCTGTCTTTGGCCGGGGC	604
176	AlaLeuThrArgValProSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGly	195
605	GAGGTGTGCCAGCTCGGCTCCAGCGTGGCAGGTGAGCGGGAAACCTCTATGCGGCC	664
196	GluValCysAspAlaArgLeuGlnArgCysGlnValSerArgGluProSerTyrAlaAla	215
665	TTCTCCACCGCGCGGCTTCCACTCGCGGTGGCGTGTGGTGTGTCTTCTACCGGAAG	724
216	PheSerThrArgGlyAlaPheHisLeuProLeuGlyValValProPheValTyrArgLys	235
725	ATCTACGAGGCGCCAAAGTTTCGTTTCGGCCGCGCGGAGAGTGTGCTGCGCTGCGG	784
236	IleTyrGluAlaAlaLysPheArgPheGlyArgArgArgAlaValLeuProLeuPro	255
785	GCCACCATTCGAGGTGAGGTCCAAAGTAAAGGAAGCACCTGATGAGGTGAAGTGGTTC	844
256	AlaThrMetGlnVal-----LysValLysGluAlaProAspGluAlaGluValValPhe	273
845	ACGGACATTTGAAAGCAACGGTGTCTTCCAGGTGACGGGGACTCCTCGCGGGACAG	904
274	ThrAlaHisCysLysAlaThrValSerPheGlnValSerGlyAspSerTrpArgGluGln	293
905	AAGGAGCGCAGCAGCAGCATGATGTTGGGAATTCGATTGGCGGTGTTGTGCTGTGCTG	964
294	LysGluArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp	313
965	ATCCCCCTTCTTCTCGACCGGAACATCATAGCCCACTCTGTGCTCGAGCTGCCCGCCATC	1024
314	IleProPhePheLeuThrGlulLeuIleSerProLeuCysAlaCysSerLeuProProIle	333
1025	TGGAAAGCATATTTCTGTGGCTGGCTACTCCAAATCTTCTTCTCAACCCCTCATTTAC	1084
334	TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr	353
1085	ACAGCTTTTAAACAAGAACTTACAACAATGCCTTCAAGAGCCTCTTTACTAAGCAAGA	1141
354	ThrAlaPheAsnLysAsnTyrAsnAlaAlaPheLysSerLeuPheThrLysGlnArg	372

```

1  APPLICANT: GORMAN, LINDA
2  APPLICANT: TCHERNEV, VELIZAR T.
3  APPLICANT: MALYANKAR, URIEL M.
4  APPLICANT: SHENOY, SURESH
5  APPLICANT: PADIGARU, MURALIDHARA
6  APPLICANT: GERLACH, VALERIE L.
7  APPLICANT: MACDOUGALL, JOHN R.
8  APPLICANT: SMITHSON, GLENDA
9  APPLICANT: MILLET, ISABELLE
10 APPLICANT: PEYMAN, JOHN
11 APPLICANT: STONE, DAVID
12 APPLICANT: GUNTHER, ERIK
13 APPLICANT: ELLERMAN, KAREN
14 APPLICANT: LI, LI
15 APPLICANT: RASTELLI, LUCA
16 APPLICANT: ZERHUSEN, BRYAN
17
18 TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
19
20 FILE REFERENCE: 21402-114
21
22 CURRENT APPLICATION NUMBER: US/09/954,342
23
24 CURRENT FILING DATE: 2001-09-17
25
26 PRIOR APPLICATION NUMBER: 60/233,382
27
28 PRIOR FILING DATE: 2000-09-18
29
30 PRIOR APPLICATION NUMBER: 60/240,498
31
32 PRIOR FILING DATE: 2000-10-13
33
34 PRIOR APPLICATION NUMBER: 60/260,284
35
36 PRIOR FILING DATE: 2001-01-08
37
38 PRIOR APPLICATION NUMBER: 60/260,973
39
40 PRIOR FILING DATE: 2001-01-11
41
42 PRIOR APPLICATION NUMBER: 60/264,794
43
44 PRIOR FILING DATE: 2001-01-29
45
46 PRIOR APPLICATION NUMBER: 60/238,398
47
48 PRIOR FILING DATE: 2000-10-06
49
50 PRIOR APPLICATION NUMBER: 60/232,675
51
52 PRIOR FILING DATE: 2000-09-15
53
54 PRIOR APPLICATION NUMBER: 60/274,862
55
56 PRIOR FILING DATE: 2001-03-09
57
58 PRIOR APPLICATION NUMBER: 60/233,801
59
60 PRIOR FILING DATE: 2000-09-19
61
62 PRIOR APPLICATION NUMBER: 60/232,676
63
64 PRIOR FILING DATE: 2000-09-15
65
66 PRIOR APPLICATION NUMBER: 60/233,960
67
68 PRIOR FILING DATE: 2000-09-20
69
70 PRIOR APPLICATION NUMBER: 60/233,402
71
72 PRIOR FILING DATE: 2000-09-18
73
74 PRIOR APPLICATION NUMBER: 60/233,521
75
76 PRIOR FILING DATE: 2000-09-19
77
78 PRIOR APPLICATION NUMBER: 60/233,522
79
80 PRIOR FILING DATE: 2000-09-19
81
82 PRIOR APPLICATION NUMBER: 60/232,679
83
83 PRIOR FILING DATE: 2000-09-15
84
84 NUMBER OF SEQ ID NOS: 104
85
86 SOFTWARE: PatentIn Ver. 2.1
87
88 SEQ ID NO 30
89
90 LENGTH: 380
91
92 TYPE: PRT
93
94 ORGANISM: Unknown Organism
95
96 FEATURE:
97
98 OTHER INFORMATION: Description of Unknown Organism: NOVX Polypeptide
99
100 US-09-954-342-30

```

RESULT 6

RESOLUTION 954-342-30
; Sequence 30, Application US/09954342
; Publication No. US20030170838A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, VISHNU S.
; APPLICANT: SPYTEK, KIMBERLY ANN
; APPLICANT: TAUPIER, RAYMOND J.
; APPLICANT: VERNET, CORINE A.
; APPLICANT: COLMAN, STEVEN D.

QY 65 ACCAGCAGC-----GGACCCGGGACC 85
Db 19 ThrSerSerProAlaLeuProLeuProTrpAspProArgProAlaalaGlyProGlyThr 38
QY 86 CCAAGCCGAGAGGATACCTGGTTCGACCCCGAGCGGCGCGTCTCGCGCGCGAGG 145
Db 39 ProSerProArgGlyLeuGlySerThrProSerGlyAlaValLeuProGlyArgGly 58
QY 146 CGCCCTCTCTGCTCTTACCGTCTCGGTGGTGAGCTGCTAGTGTGCTGATCGCTGC 205
Db 59 ProProPheSerValPheThrValLeuValValThrLeuLeuValLeuLeuAla 78
QY 206 ACTTTCCTGTGAACCTGCTGCTTCGCTACCATCCCGCGGTCCGTGCCCTTCCACCGC 265
Db 79 ThrPheLeuTrpAsnLeuValProValThrilleProArgValArgAlaPheHisArg 98
QY 266 GTGCGGCATACCTGGTGGCTCGAGCGCGCTCTCGAGAACTAGTGAGCGCTGGCG 325
Db 99 ValProHisAsnLeuValAlaSerThrAlaValSerAspGluLeuValAlaAlaLeuAla 118
QY 326 ATGCCACCGAGCCTCGCGAGTGTGCTGCACCGCGGCGACGTCGGCTGTGGCGCGGAGC 385
Db 119 MetProProSerLeuAlaSerGluLeuSerThrGlyArgArgLeuLeuGlyArg--- 137
QY 386 CTGTGCCACGTGTGGATCTCTTCCGACCGCGGAGCCTGTGTGTGCTGCGCCCGCGCCTC 445
Db 138 -----HisValTrpIleSerPheAspAla-----LeuCysCysProAlaGlyLeu 152
QY 446 GGGACGTGGCGGCATCGCTCGCGCGCGCGAGCGGCATCACAGGACCTGCGAGCAC 505
Db 153 GlyAsnValAlaAlaLeuLeuGlyArgAspGlyAlaIleThrArgHisLeuGlnHis 172
QY 506 ACGTGTGCGCACCGCGCGCGCTGTGTGCTATGATCGCGCTCGCGCGCGCGCGCTCG 565
Db 173 ThrLeuArgThrArgSerArgAlaSerLeuLeuMetIleAlaLeuAlaArgValProSer 192
QY 566 GCGTCTATCGCCCTCGCGCGCGCTGTCTTTGGCGGGCGAGGTGTGCGAGCGCTCGGCTC 625
Db 193 AlaLeuIleAlaLeuAlaProLeuLeuPheGlyArgGlyGluValCysAspAlaArgLeu 212
QY 626 CAGCGTGCAGAGTGTGAGCGCGGACCTCTATGCGCGCTTCTCCACCGCGCGCGCTTC 685
Db 213 GlnArgCysGlnValSerArgGluProSerTyrAlaAlaPheSerThrArgGlyAlaPhe 232
QY 686 CACCTGCGCTGTGCGGTGCTGCTTGTCTACCGAAGATCTACGAGCGCGCGCACTTT 745
Db 233 HisLeuProLeuGlyValValProPheValTyrArgLysIleTyrGluAlaAlaLysPhe 252
QY 746 CGTTTCGCGCGCGCGAGAGCTGTGCTGCGGTGCGCGCGCACCATGCGAGTGCAGTCC 805
Db 253 ArgPheGlyArgArgArgAlaValLeuProLeuProAlaThrMetGln----- 269
QY 806 AAGTAAAGGAGACCTGATGAGCTGAGTGTTCACGGCACATTCGAAGCAACG 865
Db 270 ---ValLysGluAlaProAspGluAlaGluValValPheThrAlaHisCysLysAlaThr 288
QY 866 GTGTCTTCAGGAGAGCGGAGCTCTGCGCGGAGCAGAGGAGCGAGCGCGCGCATG 925
Db 289 ValSerPheGlnValSerGlyAspSerTrpArgGluGlnLysGluArgAlaAlaMet 308
QY 926 ATGGTGGGAATCTGATGGCGTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 985
Db 309 MetValGlyIleLeuIleGlyValPheValLeuCysTrpIleProPhePheLeuThrGlu 328
QY 986 CTATCAGCCCACTCTGTGCTGCGAGCTGCGCCCACTGCGAAAGCATATTTCTGTGG 1045
Db 329 LeuIleSerProLeuCysAlaCysSerLeuProIleTrpLysSerIlePheLeuTrp 348
QY 1046 CTTGGCTACTCCAACTCTTCTTCAACCCCTGATTTTACAGCTTTTAAACAAGAACTAC 1105
Db 349 LeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyrThrAlaPheAsnLysAsnTyr 368

QY 1106 AACATGCTCTCAAGAGCGCTCTTTACTAAGCAGAGA 1141
Db 369 AsnAsnAlaPheLysSerLeuPheThrLysGlnArg 380
RESULT 7
US-09-823-187-56
; Sequence 56, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Patturajan, Muralidhar
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 56
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-823-187-56
Alignment Scores:
Pred. No.: 4.04e-82 Length: 370
Score: 1453.50 Matches: 296
Percent Similarity: 83.38% Conservative: 20
Best Local Similarity: 78.10% Mismatches: 54
Query Match: 67.48% Indels: 9
DB: 10 Gaps: 3
US-09-976-782-15 (1-1152) x US-09-823-187-56 (1-370)
QY 5 ATGAGCGCCCTAGCCTTTTCAGTGGCCACCGCGCGCTTGCCTTGCCTGGAGACCGCAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProProGlyProGlu 20
QY 65 ACCAGCAGCGAGCCCGGACCCCAAGCCCGAGAGGATATCTCGGTTTCGACCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCGCTCTCGCGCGCGCGCGCGCTTCTCTGTCTTTCACGGTCTCTGTGTGTGACGCTG 184
Db 39 LeuIleLeuSerGlyArgGluProProPheSerAlaPheThrValLeuValValThrLeu 58

```
QY 185 CTAGTGTGCTGATCGCTGCCACTTTCCTGTGGAACCTGCTGGTTCCTCGGTCAACATCCG 244
Db |||||||
QY 59 LeuValLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 78
Db |||||||
QY 245 CGGTCCTGCTCCATCCAGCGGTGCGGCATACCTTGGTGGCTCGACCGCGCTCTCGAC 304
Db |||||||
QY 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
Db |||||||
QY 305 GAACATAGTGGAGCGCTGCGATGCCACCGACCTGGCGAGTGAGCTGTCGACCGGGCA 364
Db |||||||
QY 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
Db |||||||
QY 365 CGTGGCTGCTGGCGCGAGCTGTGCCACGTGTGATCTCTTCGACCGCGGAGCTGT 424
Db |||||||
QY 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
Db |||||||
QY 425 CTGTGCTGCTGCCCGCGCTCGGACGTGGCGGCATCGCTCGCGCGCGACGCGGGCC 484
Db |||||||
QY 136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTrpThr 155
Db |||||||
QY 485 ATCAGCGGACCTGCGACACAGCTGCGCACCGCGCGCGCTGCTGCTCATGATC 544
Db |||||||
QY 156 IleThrArgHisLeuGlnTrpThrLeuArgTrpArgArgAlaSerAlaLeuMetIle 175
Db |||||||
QY 545 GCGCTGCGCGGTGCGGTGCGGCTCATCGCTCGCGCGCTGCTCTTTGCGCGGGC 604
Db |||||||
QY 176 AlaIleThrTrpAlaLeuSerAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 195
Db |||||||
QY 605 GAGGTGTGCGAGCTCGGCTCCAGCGCTGCCAGTGCAGCGCGGAACTCTCTATGCGGCC 664
Db |||||||
QY 196 GluAlaTyArgAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTrpAlaVal 215
Db |||||||
QY 665 TTCTCCACCGCGCGGCTTCCACCTGCGCTGCGGTGCGGTGCTGCTACCGGAAG 724
Db |||||||
QY 216 PheSerThrCysGlyAlaPheTrpValProLeuAlaValLeuPheValTrpLys 235
Db |||||||
QY 725 ATCTAGAGCGGCCAAGTTTCTGTTGCGCGCGCGCGAGAGCTGCTGCTGCGTTCGCG 784
Db |||||||
QY 236 IleTyTrpLysAlaAlaLysPheArgPheGlyArgArgArgAlaValValProLeuPro 255
Db |||||||
QY 785 GCCACCATGAGGTGAGGTCCAAAGTAAAGAAAGCACTGATGAGGTGAAGTGGTTC 844
Db |||||||
QY 256 AlaThrThrGln-----AlaLysGluAlaProGlnGluSerGluThrValPhe 271
Db |||||||
QY 845 ACGGCATTTGAAAGCAAGCGTGTCTCCAGTGCAGCGGAGCTCCTCGCGGGAGCAG 904
Db |||||||
QY 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTrpArgGluGln 291
Db |||||||
QY 905 AAGGAGAGGAGGAGCAGCATGATGGTGGGAATTCGATTGGCGTGTGTGTGCTGTGCTGG 964
Db |||||||
QY 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
Db |||||||
QY 965 ATCCCTTCTTCTGACGGAACCTATCAGCCCACTGTGTGCTGCGCTGCGCTGCGCCCATC 1024
Db |||||||
QY 312 IleProPheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProProIle 331
Db |||||||
QY 1025 TGAAGAAGCATATTTCTGTGCTGGCTGCTACCAATTTCTTCAACCCCTGATTATC 1084
Db |||||||
QY 332 TrpLysSerIlePheLeuTrpLeuGlyTyTrSerAsnSerPhePheAsnProLeuIleTy 351
Db |||||||
QY 1085 ACAGCTTTTAAACAGAACTCAACAATGCTTCAAGAGCTCTTTTAAAGCAGAGA 1141
Db |||||||
QY 352 ThrAlaPheAsnLysAsnTyAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370
Db |||||||
```

RESULT 8

```
US-09-923-187-57
; Sequence 57, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shinkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Rattus norvegicus
; US-09-823-187-57

Alignment Scores:
Pred. No.: 4,04e-82 Length: 370
Score: 1453.50 Matches: 296
Percent Similarity: 83.38% Conservative: 20
Best Local Similarity: 78.10% Mismatches: 54
Query Match: 67.48% Indels: 9
DB: 10 Gaps: 3

US-09-976-782-15 (1-1152) x US-09-823-187-57 (1-370)
QY 5 ATGAGGCGCGCTAGCCTTTTCAGTGGCCACCGCGCGGTTCCTTGCCTGGGACCCGAG 64
Db |||||||
QY 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyIleAlaPheProGlyProGlu 20
Db |||||||
QY 65 ACCAGAGCGGACCCCGGACCCCAAGCCGAGAGGATACCTCGGTTCCGACCCCGAGCGC 124
Db |||||||
QY 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
Db |||||||
QY 125 GCGCTCTCTCGCGCGGCGGCGCGCTTCTCTGTCTTCTACGGTCTCGGTGGTGGACCGCTG 184
Db |||||||
QY 39 LeuIleLeuSerGlyArgGluProProPheSerAlaPheThrValLeuValValThrLeu 58
Db |||||||
QY 185 CTAGTGTGCTGATCGCTGCCACTTTCTCTGTGGAACCTGCTGGTTCGCGTCAACATCCG 244
Db |||||||
QY 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrIleLeu 78
Db |||||||
QY 245 CGGTCCTGCTTCCACCGCGCTGCCATAACTTGGTGGCGCTCGACCGCGCTCTCGGAC 304
Db |||||||
QY 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
Db |||||||
QY 305 GAACATAGTGGAGCGCTGCGATGCCACCGACCTGGCGAGTGAGCTGTCGACCGGGCA 364
Db |||||||
QY 99 ValLeuValAlaAlaLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
Db |||||||
QY 365 CGTGGCTGCTGGCGCGGAGCTGTGCCACGTGTGATCTCTTCCGAGCGCGGAGCTGT 424
Db |||||||
```



```
QY 605 GAGGTGTGCGACGCTCGCTCCAGCGCTGCCAGGTGAGCCGGGAAACCTCCCTATCCCGCC 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnLupProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGCGCTCCACCTGCGCTTGGCGTGGTGGTGGTGGTGGTGGTGGTGGTGG 724
Db 216 PheSerThrCysGlyAlaPheTyrValProLeuAlaValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGGCGCGCAAGTTTCTGTTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
QY 785 GCCACCATGCGAGTGCAGGTCACAGGTAAGGAGCACCTGATGAGCGTGAAGTGTGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProGlnLupProGlnLupPro 271
QY 845 ACGGCACATGTCAGAACAGCGTGTCTCCAGGTGAGCGGGAGCTCTCGCGGGAGCAG 904
Db 272 ThrAlaArgCysArgAlaThrValAlaPheGlnThrSerGlyAspSerTyrArgGluGln 291
QY 905 AAGGAGGCGGAGCAGCATGATGCTGGGAATCTGATGAGGCTGTTGCTGCTGCTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCGCTTCTTCTGACGGAATCATACAGCCACTCTGCTGCTGAGCTGCGCGCGCGCG 1024
Db 312 IleProPhePheLeuThrGluLeuValSerProLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAGCATATTCTCTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1084
Db 332 TrpLysSerIlePheLeuThrLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTACAGAACTCAACAACTGCTTCAAGAGCTTCTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 11
US-09-823-187-58
; Sequence 58, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catherine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderina, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,005
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: 60/195,792
; PRIOR FILING DATE: 2000-04-10
; PRIOR APPLICATION NUMBER: 60/196,556
; PRIOR FILING DATE: 2000-04-11
; PRIOR APPLICATION NUMBER: 60/197,081
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/197,525
```

```
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/197,087
; PRIOR FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 103
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-823-187-58

Alignment Scores:
Pred. No.: 4,66e-82 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.64% Conservative: 19
Best local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 10 Gaps: 3

US-09-976-782-15 (1-1152) x US-09-823-187-58 (1-370)
```

```
QY 5 ATGAGAGCGCGTAGCCTTTCAGTGGCCACCGCGCGCTTGCCTTGCCTGGAGCCCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProProGlyProGlu 20
QY 65 ACCAGCAGCGGACCGCGGACCCCAAGCCGAGAGGAGTACTCGGTTCACCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCCGTCTGCGCGCGCGCGCGCGCTTCTCTGCTTCTCAGGCTCCTGGTGTGACGCTG 184
Db 39 LeuIleLeuProGlyArgGluProPheSerAlaPheThrValLeuValValThrIleu 58
QY 185 CTAGTGTCTGATCGCTGCGCACCTTCTGTGTGGAACCTGCTGGTTCGGGTCAACATCCG 244
Db 59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuLeuValLeuValThrIleu 78
QY 245 CGGGTCTGCTGCTTCCACCGCGTCCGCATTAACCTTGGTGGCTCGAGCGCGCTCTCGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACCTAGTGGCAGCGCTGGCGATGCCACCGAGCTGGCGAGTGTCTGACCGCGGCGA 364
Db 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTGTGGCGCGGAGCCTGTGCGACCTGTGTGGATCTCTCTGACCGCGGAGCCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCCCGCGCGCTCGGGAACGTGGCGGCCCATCGCCCTGGGCGCGAGCGGCGCC 484
Db 136 LeuCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAlaPheArgTyrTrpThr 155
QY 485 ATCACAGGCGACCTGACGACACACCTGCGCACCCCGAGCGCGCTCTGCTCATGATC 544
Db 156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCGCGGCGTGGCGCTCATCGCCTCGCGCGCTGCTCTTTGCGCGGCGGC 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGGTGTGCGAGCGCTCGGCTCCAGCGTGCAGGTGAGCGCGGAGCCCTCTATGCGCGCC 664
Db 196 GluAlaTyrAspAlaArgLeuLeuArgCysGlnValSerGlnProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCTTGGCTGGTGGCTGGTGGTGGTGGTGGTGG 724
Db 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGGCGCGCAAGTTTCTGTTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
```

```
QY 785 GCCACCATGCGAGTCCAAAGTAAAGGAAGACACCTGATGAGCGTGAAGTGTGTC 844
Db ||||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
256 AlaThrThrGln-----AlaLysGluAlaProGluSerGluMetValPhe 271
QY 845 ACGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGACTCTCGCGGAGCAG 904
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
QY 905 AAGGAGAGCGGAGCAGCCATGATGTGGGAATCTGATTTGGGTGTGTGTGCTGG 964
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTTCTCCTGAGCAACTCATCAGCCACTGTGCTGAGCGCTGCCCCCATC 1024
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProPhe 331
QY 1025 TGGAAAGCATATTTCTGTGGTGTGGCTACTCAATTTCTTCAACCCCTGATTTAC 1084
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAAGAACTACAACTAGCTTCAAGAGCCTCTTACTAAGCAGAGA 1141
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 12
US-09-954-342-57
; Sequence 57, Application US/09954342
; Publication No. US20030170838A1
; GENERAL INFORMATION:
; APPLICANT: MISHRA, VISHNU S.
; APPLICANT: SPYTEK, KIMBERLY ANN
; APPLICANT: TAUIEK, RAYMOND J.
; APPLICANT: VERNET, CORINE A.
; APPLICANT: COLMAN, STEVEN D.
; APPLICANT: GORMAN, LINDA
; APPLICANT: TCHERNEV, VELIZAR T.
; APPLICANT: MALYANKAR, URIEL M.
; APPLICANT: SHENOY, SURESH
; APPLICANT: PADIGARU, MURALIDHARA
; APPLICANT: GERLACH, VALERIE L.
; APPLICANT: MACDOUGALL, JOHN R.
; APPLICANT: SMITHSON, GLENDA
; APPLICANT: MILLET, ISABELLE
; APPLICANT: PEYMAN, JOHN
; APPLICANT: STONE, DAVID
; APPLICANT: GUNTER, ERIK
; APPLICANT: ELLERMAN, KAREN
; APPLICANT: LI, LI
; APPLICANT: RASTELLI, LUCA
; APPLICANT: ZERHUSEN, BRYAN
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES AND POLYPEPTIDES ENCODED THEREBY
; FILE REFERENCE: 21402-114
; CURRENT APPLICATION NUMBER: US/09/954,342
; CURRENT FILING DATE: 2001-09-17
; PRIOR APPLICATION NUMBER: 60/233,382
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/240,498
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/260,284
; PRIOR FILING DATE: 2001-01-08
; PRIOR APPLICATION NUMBER: 60/260,973
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/264,794
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 60/238,398
; PRIOR FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: 60/232,675
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/274,862
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/233,801
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,676
```

```
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: 60/233,960
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 60/233,402
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: 60/233,521
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/233,522
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 60/232,679
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 57
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-954-342-57

Alignment Scores:
Pred. No.: 4-66e-82 Length: 370
Score: 1452.50 Matches: 298
Percent Similarity: 83.64% Conservative: 19
Best Local Similarity: 78.63% Mismatches: 53
Query Match: 67.43% Indels: 9
DB: 10 Gaps: 3

US-09-976-782-15 (1-1152) x US-09-954-342-57 (1-370)

QY 5 ATGAGGCGCGTAGCTTTCACTGCGCACCGCGGGTGGCCCTTCCCTCGGACCCGAG 64
Db ||||| :|||:||||| ||||| ||||| ||||| ||||| |||||
1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20

QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATACCTCGGTTTCGACCCCGCGGC 124
Db ||| :||| ||| :|||:||||| ||||| ||||| ||||| |||||
21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38

QY 125 GCGTCTCTCGCGGCGGAGCGCGCCCTTCTCTGTCTTCCACGTCCTCGTGTGAGCGTG 184
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58

QY 185 CTAGTGTCTGTATGCTGCCACTTTCCTGTGTGAACCTGCTGGTTCGGTCACCATCCCG 244
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
59 LeuValLeuLeuIleAlaAlaThrPheLeuTrpAsnLeuValLeuValThrIleLeu 78

QY 245 CGGGTCCGTGCGCTTCCACCGCGTCCGCGATACTTCGTGGCTCGACGCGCGCTCGAC 304
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98

QY 305 GAACCTAGTGCAGCGCTGCGGATGCCACCGAGCCTGCGGAGTGAGTGTGACCGGCGCA 364
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118

QY 365 CGTCGCGCTCTGGCGCGGAGCGCTGTCACAGTGTGATCTCTTCACACCGCGGAGCGCTGT 424
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
119 ArgTrpGlnLeuGlyArgSerLeuCysHisValTrpIleSerPheAspVal----- 135

QY 425 CTGTGCTGCCCGCGCGCTCGGGAACGTGGCGGCATCGCCCTGGCGCGGACCGGGCC 484
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
136 LeuCysCysThrAlaSerIleTrpAsnValAlaAlaIleAlaLeuAspArgTyrTrpThr 155

QY 485 ATCACACGGCACCTGCGACACACGCTGCGCACCGCGCGCGCGCTGTGTGTCATGATC 544
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
156 IleThrArgHisLeuGlnTyrThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175

QY 545 GCGCTCGCGCGGTGCGCGGCTCATCGCCCTCGCGCGCTGCTCTTTGCGCGGGGC 604
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195

QY 605 GAGGTGTGCGACGCTCGGCTCCAGCGCTGCCAGCGCTGCCAGCGCGGGAACCTCTATGCGGCC 664
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
```


Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProfile 331
QY 1025 TGGAAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCTTCAACCCCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAACAAGTACAAACATGCTTCAAGAGCTCTTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 14

US-10-109-532A-2
; Sequence 2, Application US/10109532A
; Publication No. US20030009780A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING 5-HT5B
; TITLE OF INVENTION: SEROTONIN RECEPTOR GENE DISRUPTIONS
; FILE REFERENCE: R-601
; CURRENT APPLICATION NUMBER: US/10/109,532A
; CURRENT FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/280,553
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: US 60/342,472
; PRIOR FILING DATE: 2001-12-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-109-532A-2

Alignment Scores:
Pred. No.: 370
Score: 1452.50
Matches: 298
Percent Similarity: 83.64%
Best Local Similarity: 78.63%
Query Match: 67.43%
DB: 14
Gaps: 3

US-09-976-782-15 (1-1152) x US-10-109-532A-2 (1-370)

QY 5 ATGAGGCGCTAGCTTTCAGTGGCCACCGCGGGTGGCCCTTCCCTGGGACCCGAG 64
Db 1 MetGluValSerAsnLeuSerGlyAlaThrProGlyLeuAlaPheProGlyProGlu 20
QY 65 ACCAGCAGCGGACCCCGGACCCCAAGCCGAGAGGATATCTGGTTCGACCCCGAGCGGC 124
Db 21 SerCysSer-----AspSerProSerSerGlyArgSerMetGlySerThrProGlyGly 38
QY 125 GCGTCTCGCGGGCGAGGGCGCGCTTCTGTCTTCTACGGTCTCTGTGGTGAGCGCTG 184
Db 39 LeuIleLeuProGlyArgGluProProPheSerAlaPheThrValLeuValThrLeu 58
QY 185 CTAGTCTCTGATCGCTGCACCTTTCCTGTGAACCTGTGTCTCGGTCACCATCCCG 244
Db 59 LeuValLeuLeuIleAlaThrPheLeuTrpPAsnLeuLeuValLeuValThrLeu 78
QY 245 CGGGTCCGTCTCCACCGCGTGGCGCATPACTTGTGGCTCGACGGCGCTCTCGGAC 304
Db 79 ArgValArgAlaPheHisArgValProHisAsnLeuValAlaSerThrAlaValSerAsp 98
QY 305 GAACCTAGTGGCGGCTGGCGATGCCACCGAGCTGGCGAGTGGCTGCACCGGGCGA 364
Db 99 ValLeuValAlaValLeuValMetProLeuSerLeuValSerGluLeuSerAlaGlyArg 118
QY 365 CGTGGCTGTGGCGGAGCTGTGCACCTGTGGATCTCTTTCGACGCGCGAGCGCTGT 424
Db 119 ArgTrpGlnLeuGlyArgSerLeuLeuValTrpIleSerPheAspVal----- 135
QY 425 CTGTGCTGCCCGCGCGCTTGGGAACTGTGGCGGCATTCGCTGGCGCGCGAGCGGCC 484

Db 136 LeuCysCysThrAlaSerIleTrpAsnValAlaIleAlaLeuAspArgTyrTrpThr 155
QY 485 ATCACACGCGACCTCAGCACACAGCTGGCGCACCCGCGCGCTCGTTGCTCATGATC 544
Db 156 IleThrArgHisLeuGlnIleThrLeuArgThrArgSerArgAlaSerAlaLeuMetIle 175
QY 545 GCGCTCGCCCGGGTGCCTCGCGCTCATCGCCCTCGCGCGCTCTCTTTGGCGGGGC 604
Db 176 AlaIleThrTrpAlaLeuSerAlaLeuIleAlaLeuAlaProLeuLeuPheGlyTrpGly 195
QY 605 GAGGTGTCGAGCTCGCTCGCTCCAGCGCTGCACGGTGCAGCGGAGCCCTCTATCGCGC 664
Db 196 GluAlaTyrAspAlaArgLeuGlnArgCysGlnValSerGlnGluProSerTyrAlaVal 215
QY 665 TTCTCCACCGCGCGCGCTTCCACCTGCGCTTGGCTGGCTGCCCTTGTCTACCGGAAG 724
Db 216 PheSerThrCysGlyAlaPheTyrLeuProLeuAlaValLeuPheValTyrTrpLys 235
QY 725 ATCTACGAGCGGCGCAAGTTTCGTTTCGCGCCCGCGCGAGAGCTGTGCTGCTGCGG 784
Db 236 IleTyrLysAlaAlaLysPheArgPheGlyArgArgAlaValValProLeuPro 255
QY 785 GCCACATGCAGTGCAGGTCCAAAGCTAAAGGAGCACCTGATGAGGCTCAAGTGGTTC 844
Db 256 AlaThrThrGln-----AlaLysGluAlaProProGluSerGluMetValPhe 271
QY 845 ACGGCACATTGCAAGCAACGGTGTCTTCCAGGTGAGCGGGGACTCTCTGGCGGAGCAG 904
Db 272 ThrAlaArgArgAlaThrValThrPheGlnThrSerGlyAspSerTrpArgGluGln 291
QY 905 AAGGAGGCGGAGCGCATGATGTGGGAATTCGATTGGCGTGTGCTGTGCTGCTGG 964
Db 292 LysGluLysArgAlaAlaMetMetValGlyIleLeuIleGlyValPheValLeuCysTrp 311
QY 965 ATCCCTCTTCTCTGACGAACTCATCAGCCACTGTGCTGCGAGCTGCGCGCTGCCCCATC 1024
Db 312 IleProPhePheLeuThrGluLeuIleSerProLeuCysAlaCysSerLeuProProfile 331
QY 1025 TGGAAAAGCATATTTCTGTGGCTTGGCTACTCCAAATTTCTTCTTCAACCCCTGATTAC 1084
Db 332 TrpLysSerIlePheLeuTrpLeuGlyTyrSerAsnSerPhePheAsnProLeuIleTyr 351
QY 1085 ACAGCTTTTAAACAAGTACAAACATGCTTCAAGAGCTCTTTACTAAGCAGAGA 1141
Db 352 ThrAlaPheAsnLysAsnTyrAsnAsnAlaPheLysSerLeuPheThrLysGlnArg 370

RESULT 15

US-09-823-187-59
; Sequence 59, Application US/09823187
; Publication No. US20030096952A1
; GENERAL INFORMATION:
; APPLICANT: Burgess, Catharine
; APPLICANT: Gusev, Vladimir Y
; APPLICANT: Liu, Xiaohong
; APPLICANT: Majumder, Kumud
; APPLICANT: Padigaru, Muralidhar
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkets, Richard A
; APPLICANT: Spaderna, Steven K
; APPLICANT: Spytek, Kimberly
; APPLICANT: Taupier, Raymond J
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-745
; CURRENT APPLICATION NUMBER: US/09/823,187
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/193,339
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/193,205
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: 60/195,343
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: 60/195,088
; PRIOR FILING DATE: 2000-04-06

This Page Blank (uspto)